Studies in Computational Intelligence 849

Roger Lee *Editor*

Computer and Information Science

Springer

Studies in Computational Intelligence

Volume 849

Series Editor

Janusz Kacprzyk, Polish Academy of Sciences, Warsaw, Poland

The series "Studies in Computational Intelligence" (SCI) publishes new developments and advances in the various areas of computational intelligence—quickly and with a high quality. The intent is to cover the theory, applications, and design methods of computational intelligence, as embedded in the fields of engineering, computer science, physics and life sciences, as well as the methodologies behind them. The series contains monographs, lecture notes and edited volumes in computational intelligence spanning the areas of neural networks, connectionist systems, genetic algorithms, evolutionary computation, artificial intelligence, cellular automata, self-organizing systems, soft computing, fuzzy systems, and hybrid intelligent systems. Of particular value to both the contributors and the readership are the short publication timeframe and the world-wide distribution, which enable both wide and rapid dissemination of research output.

The books of this series are submitted to indexing to Web of Science, EI-Compendex, DBLP, SCOPUS, Google Scholar and Springerlink.

More information about this series at http://www.springer.com/series/7092

Roger Lee
Editor

Computer and Information Science

Springer

Editor
Roger Lee
Software Engineering and Information
Technology Institute
Central Michigan University
Mount Pleasant, MI, USA

ISSN 1860-949X ISSN 1860-9503 (electronic)
Studies in Computational Intelligence
ISBN 978-3-030-25212-0 ISBN 978-3-030-25213-7 (eBook)
https://doi.org/10.1007/978-3-030-25213-7

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Foreword

The purpose of the 18th IEEE/ACIS International Conference on Computer and Information Science (ICIS 2019) held on June 17–19, 2019 in Beijing, China was to together researchers, scientists, engineers, industry practitioners, and students to discuss, encourage, and exchange new ideas, research results, and experiences on all aspects of Applied Computers and Information Technology, and to discuss the practical challenges encountered along the way and the solutions adopted to solve them. The conference organizers have selected the best 13 papers from those papers accepted for presentation at the conference in order to publish them in this volume. The papers were chosen based on review scores submitted by members of the program committee and underwent further rigorous rounds of review.

In chapter "Feature Representation and Feature Matching for Heterogeneous Defect Prediction", Thae Hsu Hsu Mon and Hnin Min Oo propose a framework that aims to exploit cost-sensitive principal component analysis (PCA) and the feature matching to build highly effective prediction model for Software Defect Prediction (SDP).

In chapter "Learning Neural Circuit by AC Operation and Frequency Signal Output", Masashi Kawaguchi, Naohiro Ishii and Masayoshi Umeno used analog electronic circuits using alternative current to realize the neural network learning model. This model suggests the realization of the deep learning model regarding the proposed analog hardware neural circuit.

In chapter "Preprocessing with Contrast Enhancement Methods in Bone Age Assessment", Aye Aye Aung and Zin Mar Win tested contrast enhancement techniques such as Contrast Limited Adaptive Histogram Equalization (CLAHE), Histogram Equalization (HE), and Power Law Transform (PLT) techniques to improve the robustness and accuracy of Bone age assessment (BAA).

In chapter "Mis.Config: Finding Misreferred Configuration Bugs in Web Application Using Thin Slicing", Minami Yoda, Yuichi Sei, Yasuyuki Tahara and Akihiko Ohsuga propose a bug-finding tool called Mis.Config using static analysis. In their experiment, we applied our tool to real-world software to investigate whether Mis.Config can find misreferenced configurations.

In chapter "Self-Adaptation for Heterogeneous Client-Server Online Games", Satoru Yamagata, Hiroyuki Nakagawa, Yuichi Sei, Yasuyuki Tahara and Akihiko Ohsuga apply a MAPE loop model, which consists of four key activities (monitoring, analysis, planning, and execution) for adaptation, to reduce the time lag problem in a heterogeneous online game environment.

In chapter "A Method for Sharing Cell State for LSTM-Based Language Model", Seongik Park and Yanggon Kim propose a method for sharing cell state for a neural network-based language model, which considers all preceding cell states as a cell-stack. Their model achieved better performance compared to a traditional LSTM-based language model improving average perplexity scores.

In chapter "Emotional Analysis with News Using Text Mining for Framing Theory", Jinhyuck Choi, Kwangmi Ko Kim and Yanggon Kim aim to extend framing research to a different level by using computer-assisted analysis. This new method allows them to analyze massive data and to visualize the representation through text mining, natural language processing, and emotion lexicon. Within the broad framing approach, this study intends to show how major newspapers of several countries depict leaders of North Korea and South Korea.

In chapter "A Travel Decision Support Algorithm: Landmark Activity Extraction from Japanese Travel Comments", Siya Bao, Masao Yanagisawa and Nozomu Togawa analyze the feasibility of exploring landmark activity queries and representative examples from travel comments. Contributions in this paper include a framework for extracting activity concerned keywords and queries, quantifying the relationship between landmark activity and comment contents.

In chapter "Design of Intelligent Artificial Agents: Its Application in Joint Attention Task for Children with Autism", Vishav Jyoti and Uttama Lahiri designed realistic intelligent artificial agents with Indian look and capable of behaving naturally for computer-based applications for skill training for children with Autism.

In chapter "Automated Assessment of ER Model Using the Domain Knowledge", Muhammad Javed and Yuqing Lin propose an automated assessment approach, which focuses on some major issues of Entity Relationship Model (ERM) such as completeness and correctness. The results show the proposed approach is having a noticeable improvement over the existing approaches.

In chapter "Analysis of Online Social Network after an Event", Myat Mon Oo and May Thu Lwin apply Social Network Analysis (SNA) to extract knowledge from social media data. The result of this study will help to understand the structure of the network and also detect the most influential users.

In chapter "Hybrid Radius Particle Swarm Optimization Applications", Mudarmeen Munlin designed and investigated different approaches of Hybrid Radius Particle Swarm Optimization (HRPSO) algorithm, such as adaptive mutation, forward–backward propagation, and k-means combined with the Radius Particle Swarm Optimization (RPSO) to solve several real-world optimization problems. The efficiency of the proposed method is tested against the existing methods. The results show that the HRPSO gives better optimum results.

In chapter "Classification of Cell Nuclei Using CNN Features", Yuji Iwahori, Yuya Tsukada, Takashi Iwamoto, Kenji Funahashi, Jun Ueda and M. K. Bhuyan propose a method to discriminate benign or malignant of cell image. Proposed method uses a CNN to extract features from the nuclei images detected from the original cell images and benign or malignant is automatically classified by two classes of classification with SVM. The effectiveness of the proposed method has been confirmed by experiments from the viewpoint of cell classification of benign or malignant.

In chapter "An Exploratory Study on the Benefit Sharing of Broadcasting Content Industry", Seung-Hyeok Baek, Min-Gwon Park, Jin-Hua Zhang and Jun-Ho Lee examine the benefit sharing between organizations as a way to strengthen competitiveness in terms of co-evolution through inter-organizational relations in the Korea's broadcasting contents industry.

In chapter "A Study of Utilizing Backend as a Service (BaaS) Space for Mobile Applications", Cheong Ghil Kim explains how Parse Server is interoperable with an open-source based game and a single-board computer to develop use cases for the purpose of IoT programming education. This research aims to get insights of methodology for IoT programming with implementation of game ranking service and remote sensing systems. Implementation results were confirmed with Parse Dashboard and Rasberry Pi.

It is our sincere hope that this volume provides stimulation and inspiration, and that it will be used as a foundation for works to come.

Sault Ste. Marie, Canada
June 2019

Simon Xu
Algoma University

Contents

Contributors

Aye Aye Aung Image Processing Lab, University of Computer Studies, Mandalay, Myanmar

Seung-Hyeok Baek Policy Division Future Policy Team, Korea Creative Content Agency, Naju-si, Korea

Siya Bao Department of Computer Science and Communications Engineering, Waseda University, Tokyo, Japan

M. K. Bhuyan Indian Institute of Technology Guwahati, Guwahati, India

Jinhyuck Choi Department of Computer and Information Sciences, Towson University, Towson, MD, USA

Kenji Funahashi Department of Computer Science, Nagoya Institute of Technology, Nagoya, Japan

Hnin Min Oo Software Engineering Lab, University of Computer Studies, Mandalay, Myanmar

Naohiro Ishii Department of Information Science, Aichi Institute of Technology, Toyota, Japan

Yuji Iwahori Department of Computer Science, Chubu University, Kasugai, Japan

Takashi Iwamoto College of Life and Health Sciences, Chubu University, Kasugai, Japan

Muhammad Javed School of Electrical Engineering and Computing, The University of Newcastle, Newcastle, Australia

Vishav Jyoti Centre for Cognitive and Brain Sciences, IIT Gandhinagar, Gandhinagar, Gujarat, India

Masashi Kawaguchi Department of Electrical and Electronic Engineering, Suzuka National College of Technology, Suzuka Mie, Japan

Cheong Ghil Kim Department of Computer Science, Namseoul University, Cheonan, South Korea

Kwangmi Ko Kim Department of Mass Communication, Towson University, Towson, MD, USA

Yanggon Kim Department of Computer and Information Sciences, Towson University, Towson, MD, USA

Uttama Lahiri Department of Electrical Engineering, IIT Gandhinagar, Gandhinagar, Gujarat, India

Jun-Ho Lee Department of Business Administration, Hoseo University, Cheonan, Chungnam, Korea

Yuqing Lin School of Electrical Engineering and Computing, The University of Newcastle, Newcastle, Australia

May Thu Lwin University of Computer Studies, Mandalay, Myanmar

Mudarmeen Munlin Faculty of Information Science and Technology, Mahanakorn University of Technology, Bangkok, Thailand

Myat Mon Oo University of Computer Studies, Mandalay, Myanmar

Hiroyuki Nakagawa Osaka University, Osaka, Japan

Akihiko Ohsuga University of Electro-Communications, Tokyo, Japan

Min-Gwon Park Department of Business Administration, Hongik University, Seoul, Korea

Seongik Park Department of Computer and Information Sciences, Towson University, Towson, MD, USA

Yuichi Sei University of Electro-Communications, Tokyo, Japan

Yasuyuki Tahara University of Electro-Communications, Tokyo, Japan

Thae Hsu Hsu Mon Software Engineering Lab, University of Computer Studies, Mandalay, Myanmar

Nozomu Togawa Department of Computer Science and Communications Engineering, Waseda University, Tokyo, Japan

Yuya Tsukada Department of Computer Science, Nagoya Institute of Technology, Nagoya, Japan

Jun Ueda Asahikawa Medical University, Asahikawa, Hokkaido, Japan

Masayoshi Umeno Department of Electronic Engineering, Chubu University, Kasugai, Aichi, Japan

Satoru Yamagata University of Electro-Communications, Tokyo, Japan

Masao Yanagisawa Department of Computer Science and Communications Engineering, Waseda University, Tokyo, Japan

Minami Yoda University of Electro-Communications, Tokyo, Japan

Jin-Hua Zhang Department of Business Administration, Hoseo University, Cheonan, Chungnam, Korea

Zin Mar Win Image Processing Lab, University of Computer Studies, Mandalay, Myanmar

Feature Representation and Feature Matching for Heterogeneous Defect Prediction

Thae Hsu Hsu Mon and Hnin Min Oo

Abstract Software Defect Prediction (SDP) is one of the highly influential software engineering research topics. Early within-project defect prediction (WPDP) used intra-project data. However, it has limitations in prediction efficiency for new projects and projects without adequate training data. Studies of prediction have been carried out on cross-project defect prediction models (CPDP), i.e. models that are trained using other projects historical data. Heterogeneous defect prediction (HDP) is the special case of CPDP with different metric sets of source and target project. Despite the effectiveness of existing HDP methods, they can be affected by the issue of class imbalance that may decrease prediction performance. The proposed framework aims to exploit cost-sensitive principal component analysis (PCA) and the feature matching to build highly effective prediction model.

1 Introduction

In software quality assurance (SQA), prediction of software defects is critical and one of the most interesting research themes in academic and commercial communities [1–11]. In within project defect prediction, SDP models are created based on existing intradata to evaluate the vulnerability of new software components (instances) [12–15].

The NASA and PROMISE repositories contain many open source defect datasets (http://openscience.us/repo/). WPDP works well when there are enough data suitable to train a model of defect prediction. However, where historical data are limited, performing WPDP is difficult for a new project. Cross-project Defect Prediction (CPDP) is a practical way to solve WPDP problems. It develops a predictive model utilizing one or more of a source companys projects and then applies to the target

Thae Hsu Hsu Mon (✉) · Hnin Min Oo
Software Engineering Lab, University of Computer Studies, Mandalay, Myanmar
e-mail: thaehsuhsumon@ucsm.edu.mm

Hnin Min Oo
e-mail: hninminoo@ucsm.edu.mm

R. Lee (ed.), *Computer and Information Science*, Studies in Computational Intelligence 849, https://doi.org/10.1007/978-3-030-25213-7_1

company model [16, 17]. Without historical data, it becomes possible to use these public data sets to predict defects in new projects.

Several approaches to CPDP have been developed in recent years [8, 17–23]. Most of CPDP approaches have a serious limitation: CPDP techniques require the equivalent/similar feature sets for source and target project cases. Some CPDP research consider only common features when the dataset of source and target are heterogeneous [17, 18]. However, predicting projects with various different metrics can be challenging.

CPDP is known as heterogeneous defect prediction (HDP) in the situation where different metric sets of different projects need to match. To fix the heterogeneous metric, many heterogeneous defect prediction models are developed [16, 24, 25].

By using source and target project instructive metrics, these HDP techniques are learning defect prediction models, and all are achieving highly promising outcomes.

Although existing HDP methods have the effectiveness, they suffer from the class imbalance problem which may decrease the performance of prediction. The proposed system aims to exploit feature representation by cost-sensitive principal component analysis (PCA) and the feature matching to build highly effective prediction model.

The rest of this paper is arranged as follows. Section 2 summaries the related works. The proposed model is discussed in Sect. 3. Section 4 demonstrates the experiments related to the proposed system. Finally, this paper concludes in Sect. 5 with our findings and ideas for future work.

2 Background and Related Work

In recent years, a variety of technologies for predicting defects have emerged. Prediction of software defects uses historical data from reported (or corrected) failures to predict unknown code faults. However, there is not always sufficient historical training data available from the same project or in practice it is difficult to collect. In the last few years, researchers have proposed CPDP as an alternative solution to this issue [11, 17, 27].

Zimmermann et al. [11] assessed the ability of crossproject predictions based on large-scale experiments for 12 open source projects. It indicated that in most cases, current CPDP models are not performing well. Related to the work of Zimmermann, by focusing on training data selection, He et al. [27] proposed CPDP. They noted that the performance of the prediction was related to the distributional attributes.

To select similar data from the source project, Turhan et al. [17] proposed a Nearest-Neighbor filter method. They used only nearest neighbors to build training set for each test data, which have similar metrics to local data.

In contrast to the previous work, selecting similar training data from the test data, Ma et al. [18] introduced Transfer Naive Bayes (TNB) using the specific information of all the appropriate training data features. The TNB transferred data from cross-projects into the training data weights. The prediction model was constructed on the these weighted data.

These CPDP approaches, however, are predicated on the idea that the source and target data have the similar set of software features. When the metrics features for source and target projects are distinct, these methods will not be usable.

2.1 Heterogeneous Defect Prediction

Many HDP approaches [16, 24, 25] are proposed to solve the drawbacks of CPDP. Cheng et al. [26] described HDP model (CCT-SVM) by exploiting correlation subspace.

Jing et al. [16] approved the CCA+ using unified metric representation and transfer learning depending on CCA. By learning a pair of projective transformations that enhance the correlation between source and target data, CCA+ can create the target data distribution closely related to the source.

In performing metric selection and matching techniques, HDP method (HDP-KS) was introduced by Nam and Kim [24]. First, the technique of metric selection was used to eliminate unnecessary and irrelevant data from the source. And then, the target and source metrics are matched according to their distribution similarity or correlation. With the extracted metric sets, the HDP model is developed for predicting labels of target data.

Although existing HDP techniques are effective, there is still a problem to account the class imbalance issue.

2.2 Class Imbalance Problem

Learning with class imbalance problem occurs regularly in many domain applications. Imbalanced class distribution in datasets occurs when there is insufficient representation of one class, often the one that is of greater interest, i.e. the positive or minority class. In a simpler term, this means that the number of samples of positive (minority) class is much smaller than the number of samples of negative (majority) class [28].

Most of software defect data sets are imbalanced, which means the amount of defective samples is much smaller than the amount of non-defective samples. This can lead to poor performance of prediction. That is, The likelihood of defects prediction may be low while the overall performance is high. Without taking this issue into account, the effectiveness of software defect prediction in many real-world tasks would be greatly reduced.

To address HDP class imbalance issue, Cheng et al. [26] described a cost—sensitive transfer of vector support correlation method based on CCA+. By using cost factors to reduce the effects of imbalanced data in the Support Vector Machine

(SVM) model, they took into account different misclassification costs. However, it can be applied only to Support Vector Machine (SVM) classifiers and other classifiers are not applicable.

3 Methodology

In this section, we will describe our method for HDP, that consists of two parts: feature representation and feature matching.

3.1 Feature Representation

Transfer of feature representation reproduce instance data for model building. Feature learning or representation learning is a collection of strategies that enable a system to automatically explore the representations needed to detect feature or classify raw data. Its purpose is to align features from high dimensions to low dimensions. The overview of our approach is shown in Fig. 1.

3.1.1 Pricipal Component Analysis

PCA is one of the techniques used for dimensional reductions. By creating new, artificial variables called main components, PCA achieves dimensional reduction. Each major component is a linear combination of the variables observed. There are many advantages of PCA such as handling more than two parameters: it is insensitive to scale of original data; it combines correlated parameters into single variable and it identifies variables by importance. The PCA identifies patterns in the data set and finds their differences and similarities between each attribute. It acts as a powerful model for data analysis.

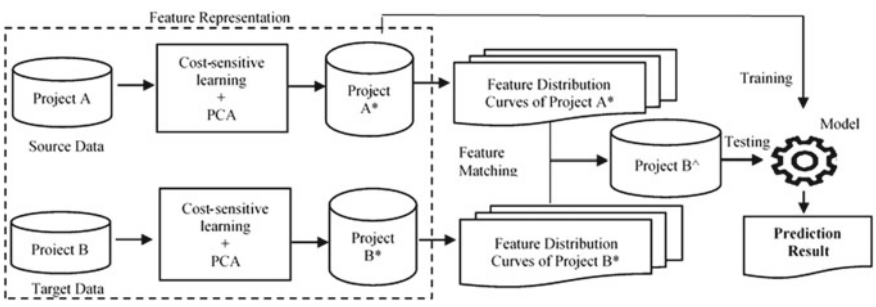


Fig. 1 Heterogenous defect prediction model framework

In the first step of PCA, the covariance matrix of the data matrix (X) is calculated. And secondly, the eigenvalues and eigenvectors for the covariance matrix are computed.

Given a data matrix ($X = [x_1; x_2; \ldots; x_N]$) , where N means the total number of samples and x_i defines the ith sample. The mean of each sample is calculated as follows.

$$\mu = \frac{1}{N} \sum_{i=1}^{N} x_i \tag{1}$$

Then, subtract the mean from all samples by the following equation:

$$D = \{d_1, d_2, \ldots, d_N\} = \sum_{i=1}^{N} x_i - \mu \tag{2}$$

After that, the covariance matrix is calculated to determine the relationship between two classes in which zero value indicates no relationship between two dimensions. Covariance matrix is achieved according to Eq. (3).

$$\Sigma = \frac{1}{N} D * D^T \tag{3}$$

Finally, the Eigenvalues and Eigenvectors for the covariance matrix are calculated by using the Eq. (4). The Eigenvalues are sorted in ascending order to select the most important data and discard the least significant data. This reduces data with higher dimensions to lower dimensions.

$$Det(A - \lambda I) = 0 \tag{4}$$

3.1.2 Cost-Sensitive Pricipal Component Analysis

Costsensitive PCA technique is used to improve the calculations of the main components, taking into account the imbalanced class issue [29]. The positive and negative samples are represented by the c^+ and c^- imbalance cost ratio respectively. In terms of mathematics,

$$C = \begin{cases} c_i^- = \frac{1-\gamma}{N^-} \\ c_i^+ = \frac{1+\gamma}{N^+} \end{cases} \tag{5}$$

In Eq. (5), for $i = 1, \ldots, N$ where, c_i^- is imbalance cost ratio for negative (non-defective) ith sample, c_i^+ equals imbalance cost ratio for positive (defective) ith sample, N^- is total number of negative sample, N^+ defines that of positive sample,

Table 1 Features in NASA dataset

No.	Feature	No.	Feature
1	Loc-blank	21	Halstead-difficulty
2	Branch-count	22	Halstead-effort
3	Call-pairs	23	Halstead-error-est
4	Loc-code-and-comments	24	Halstead-length
5	Loc-comments	25	Halstead-level
6	Condition-count	26	Halstead-prog-time
7	Cyclomatic-complexity	27	Halstead-volume
8	Cyclomatic-density	28	Maintenance-severity
9	Decision-count	29	Modified-condition-count
10	Decision-density	30	Multiple-condition-count
11	Design-complexity	31	Node-count
12	Design-density	32	Normalized-cylomatic-complexity
13	Edge-count	33	Num-openrands
14	Essential-complexity	34	Num-operators
15	Essential-density	35	Num-unique-operands
16	Loc-executable	36	Num-unique-operators
17	Parameter-count	37	Number-of-lines
18	Global-data-complexity	38	Percent-comments
19	Global-data-density	39	Loc-total
20	Halstead-content		

γ means the ratio of the size of two classes ($0 \leq \gamma \leq 1$) and y_i represents class label for each instance i. Eq. (6) represents for the weight of first principal component.

$$w_1 = \sum_{i:y_i=-1,j} (c_i - X_{ij} w_{1j}) + \sum_{i:y_i=+1,j} (c_i + X_{ij} w_{1j}) \tag{6}$$

As shown in Fig. 1, two datasets A and B are heterogeneous in nature. By applying cost-sensitive PCA, we represent features of source project A and target projects B. And then, A^* and B^* are achieved for feature matching.

3.2 Feature Matching

In the feature matching process, we calculate the following steps:

1. Obtain the all features of source A^* and target project B^*.
2. Select N samples from both projects for feature matching.

Table 2 Features in PROMISE dataset

No.	Feature	No.	Feature
1	Weighted methods per class	11	Lines of code
2	Depth of inheritance tree	12	Data access metric
3	Number of children	13	Measure of aggregation
4	Coupling between object classes	14	Measure of functional abstraction
5	Response for a class	15	Cohesion among methods of class
6	Lack of cohesion in methods	16	Inheritance coupling
7	Afferent couplings	17	Coupling between methods
8	Efferent couplings	18	Average method complexity
9	Number of public methods	19	Max-McCabes cyclomatic complexity
10	Lack of cohesion in methods	20	Avg-McCabes cyclomatic complexity

Table 3 Detail list of NASA and PROMISE datasets

Project	# of features	# of total samples	# of defective samples	# of non-defective samples	Imbalance ratio
CM1	37	327	42	285	7
MC1	38	1988	46	1942	42
MW1	37	253	27	226	8
PC1	37	705	61	644	11
PC2	36	745	16	729	46
PC3	37	1077	134	943	7
Ant1.7	20	745	166	579	3
Xalan2.4	20	723	110	613	6
Synapse1.2	20	256	86	170	2
Ivy2.0	20	352	40	312	8

3. Obtain the feature distribution curves of both projects where their values for each feature are ranked by increasing order.
4. Determine the distance and match the attributes between the different distribution curves.
5. Achieve project $\hat{B}$ that matches target project A^* features.

Project A^* is considered the set of training and Project $\hat{B}$ is considered the set of tests.

4 Experiments

4.1 Experiment Datasets

4.1.1 NASA

It is open to the public and is widely used for defect prediction. These datasets contain static code metrics including size, readability, and complexity metrics as shown in Table 1.

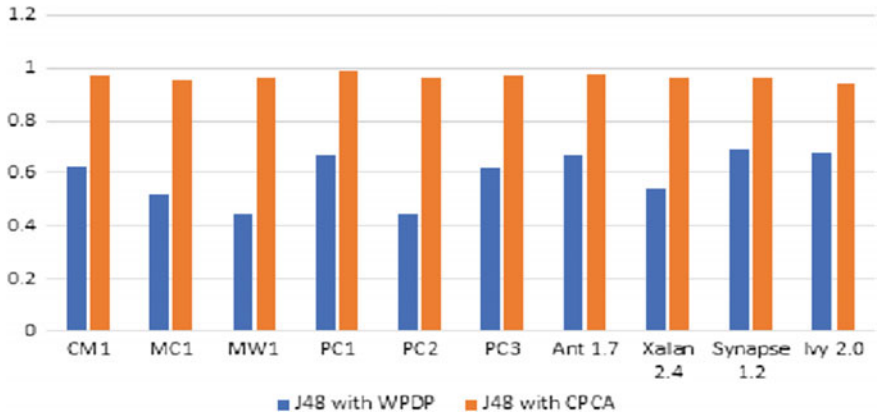


Fig. 2 Average ROC results on 10 datasets with J48 classifier

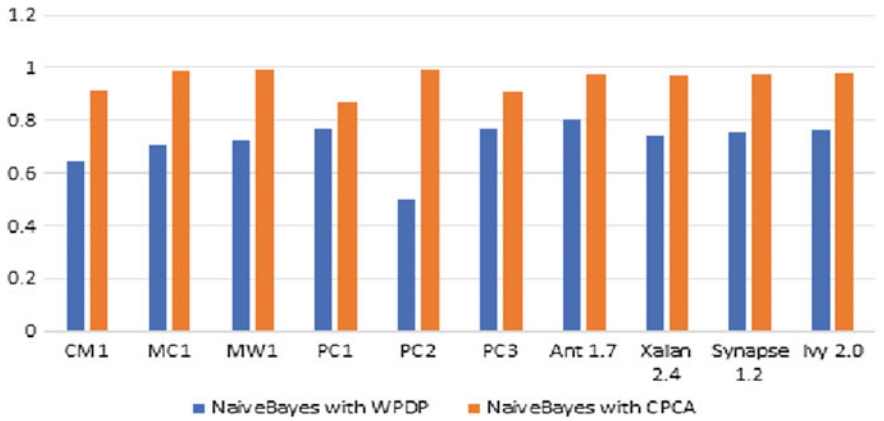


Fig. 3 Average ROC results on 10 datasets with Naive Bayes classifier

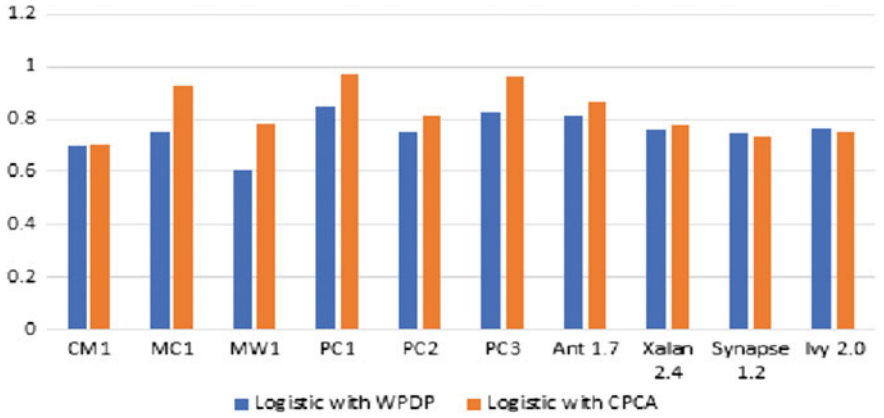


Fig. 4 Average ROC results on 10 datasets with logistic classifier

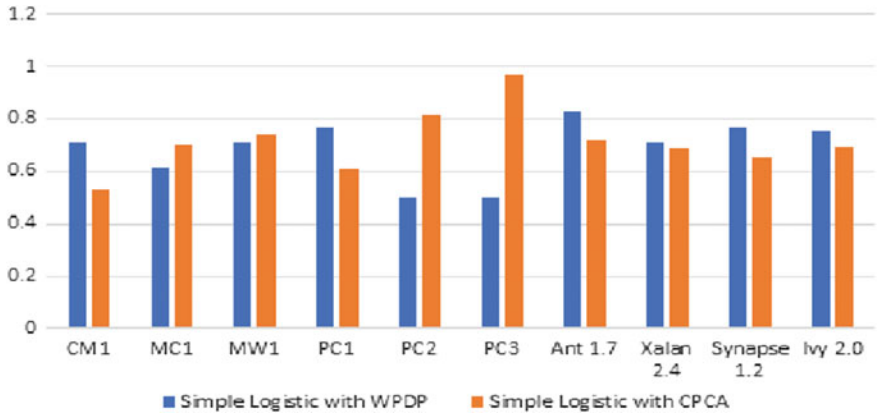


Fig. 5 Average ROC results on 10 datasets with simple logistic classifier

4.1.2 PROMISE

It is collected by Jureczko and Madeyski, consisting of many software projects, from the online data repository. It contains CK metrics, OO (Object-Oriented) metrics and McCabes cyclomatic metrics. Table 2 describes the features of PROMISE datasets.

Table 3 lists the details of these datasets.

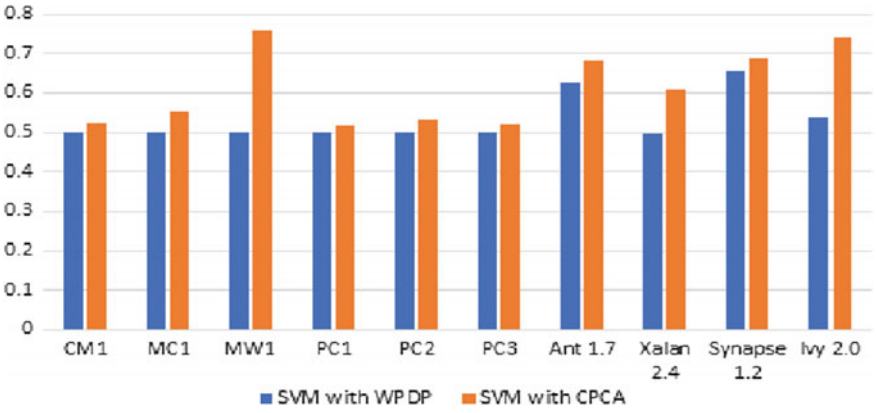


Fig. 6 Average ROC results on 10 datasets with SVM classifier

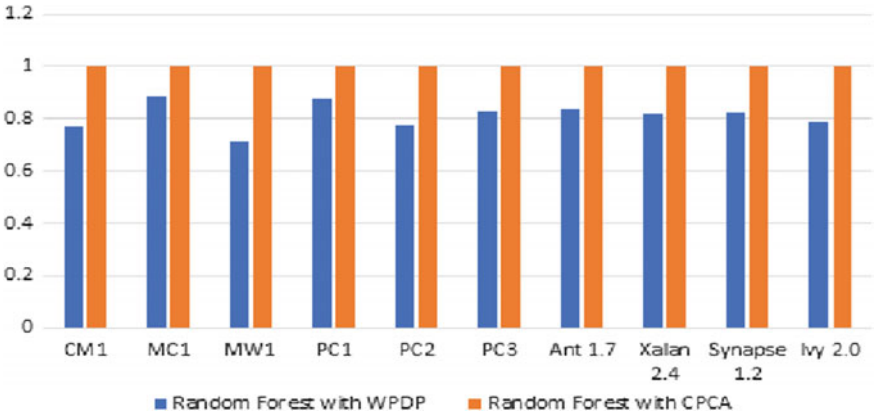
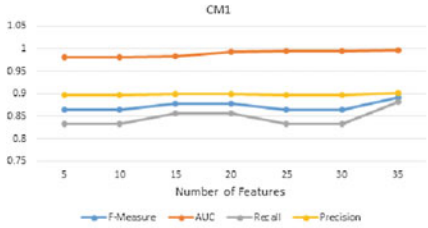


Fig. 7 Average ROC results on 10 datasets with random forest classifier

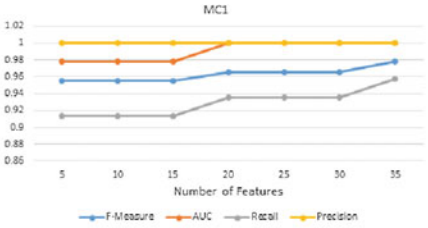
4.2 Performance Measures

To evaluate the performance of defect prediction model, Area Under the Receiver Operating Characteristic Curve (AUC) will be used. The greater AUC means the better result of prediction.

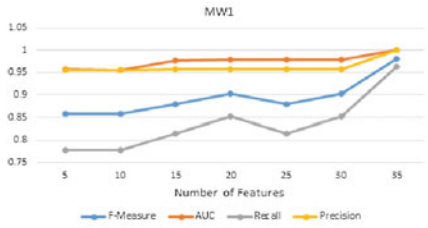
Figures 2, 3, 4, 5, 6 and 7 reports the average ROC (AUC) values of our feature representation results with comparing using cost sensitive principal component analysis and within-project defect prediction. We use six classifiers namely, J48 classifier, Naive Bayes classifier, Logistic classifier, Simple Logistic classifier, Support Vector Machine (SVM) classifier and Random Forest classifier. According to these results, our contribution feature representation can achieve the best representation and prediction results in Random Forest classifier.



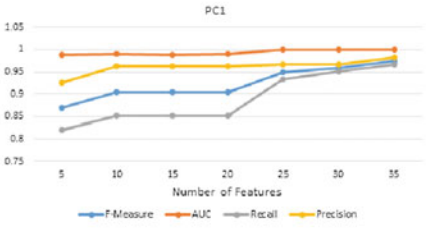
(a) Prediction results with different feature dimensionalities on CM1 project



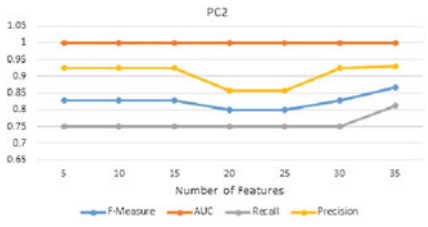
(b) Prediction results with different feature dimensionalities on MC1 project



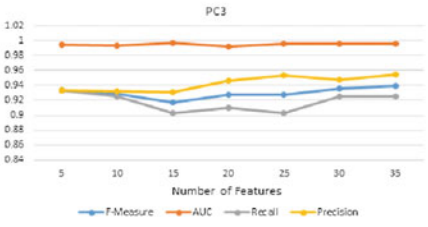
(c) Prediction results with different feature dimensionalities on MW1 project



(d) Prediction results with different feature dimensionalities on PC1 project



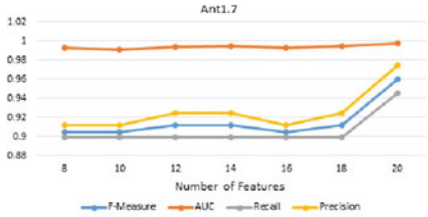
(e) Prediction results with different feature dimensionalities on PC2 project



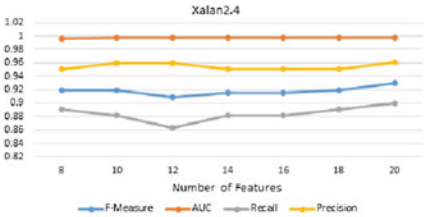
(f) Prediction results with different feature dimensionalities on PC3 project

Fig. 8 Prediction results on NASA dataset

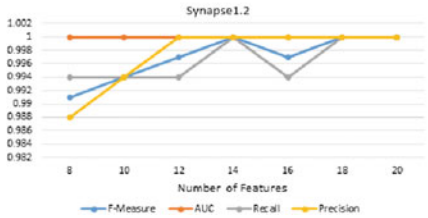
In Figs. 8 and 9, we do experiments on 6 projects of NASA datasets and 4 projects of PROMISE datasets with different number of feature dimentionalities, respectively. We use four prediction measurements: F measure, AUC, Recall and Precision by applying Random Forest classifer. A good prediction model will have high F-measure and AUC. Precision and recall both indicate accuracy of the model. Precision means the percentage of results that are relevant. Recall is the measure of defective instances that are correctly predicted as the defective samples. The higher the Recall the better the prediction performance of this approach.



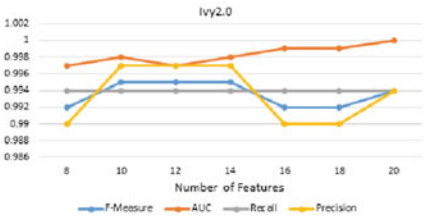
(a) Prediction results with different feature dimensionalities on Ant1.7 project



(b) Prediction results with different feature dimensionalities on Xalan2.4 project



(c) Prediction results with different feature dimensionalities on Synapse1.2 project



(d) Prediction results with different feature dimensionalities on Ivy2.0 project

Fig. 9 Prediction results on PROMISE dataset

5 Conclusion

In order to improve the software quality, heterogeneous defect prediction is very encouraging when the source and target projects have unrelated or different metric setting. It may also be used to predict the defect of projects that are in the absence of required historical data or to apply on new projects. The proposed system exploits cost-sensitive principal component analysis to eliminate the class imbalance problem defect prediction and feature matching to achieve similar distribution between source and target project. For future work, we want to use different data to validate our approachs generalization capacity involving open source projects as well as proprietary closed projects.

References

1. D'Ambros, M., Lanza, M., Robbes, R.: Evaluating defect prediction approaches: a benchmark and an extensive comparison. Empir. Softw. Eng. **17**(4–5), 531–577 (2012)
2. Fukushima, T., Kamei, Y., McIntosh, S., Yamashita, K., Ubayashi, N.: An empirical study of just-in-time defect prediction using cross-project models. In: Proceedings of the 11th Working Conference on Mining Software Repositories, pp. 172–181. ACM (2014)
3. Lee, T., Nam, J., Han, D., Kim, S., In, H.P.: Micro interaction metrics for defect prediction. In: Proceedings of the 19th ACM SIGSOFT symposium and the 13th European conference on Foundations of software engineering, pp. 311–321. ACM (2011)

4. Xu, Z., Li, S., Tang, Y., Luo, X., Zhang, T., Liu, J., Xu, J.: Cross version defect prediction with representative data via sparse subset selection. In: Proceedings of the 26th Conference on Program Comprehension, pp. 132–143. ACM (2018)
5. Lessmann, S., Baesens, B., Mues, C., Pietsch, S.: Benchmarking classification models for software defect prediction: a proposed framework and novel findings. IEEE Trans. Softw. Eng. **34**(4), 485–496 (2008)
6. Li, M., Zhang, H., Wu, R., Zhou, Z.H.: Sample-based software defect prediction with active and semi-supervised learning. Autom. Softw. Eng. **19**(2), 201–230 (2012)
7. Menzies, T., Greenwald, J., Frank, A.: Data mining static code attributes to learn defect predictors. IEEE Trans. Softw. Eng. **33**(1), 2–13 (2007)
8. Nam, J., Pan, S.J., Kim, S.: Transfer defect learning. In: 2013 35th International Conference on Software Engineering (ICSE), pp. 382–391. IEEE (2013)
9. Rahman, F., Posnett, D., Devanbu, P.: Recalling the imprecision of cross-project defect prediction. In: Proceedings of the ACM SIGSOFT 20th International Symposium on the Foundations of Software Engineering, p. 61. ACM (2012)
10. Shivaji, S., Whitehead, E.J., Akella, R., Kim, S.: Reducing features to improve code change-based bug prediction. IEEE Trans. Softw. Eng. **39**(4), 552–569 (2013)
11. Zimmermann, T., Nagappan, N., Gall, H., Giger, E., Murphy, B.: Cross-project defect prediction: a large scale experiment on data versus domain versus process. In: Proceedings of the the 7th Joint Meeting of the European Software Engineering Conference and the ACM SIGSOFT Symposium on the Foundations of Software Engineering, pp. 91–100. ACM (2009)
12. Kim, S., Zhang, H., Wu, R., Gong, L.: Dealing with noise in defect prediction. In: 2011 33rd International Conference on Software Engineering (ICSE), pp. 481–490. IEEE (2011)
13. Lee, T., Nam, J., Han, D., Kim, S., In, H.P.: Developer microinteraction metrics for software defect prediction. IEEE Trans. Softw. Eng. **42**(11), 1015–1035 (2016)
14. Jing, X.Y., Ying, S., Zhang, Z.W., Wu, S.S., Liu, J.: Dictionary learning based software defect prediction. In: Proceedings of the 36th International Conference on Software Engineering, pp. 414–423. ACM (2014)
15. Liu, M., Miao, L., Zhang, D.: Two-stage cost-sensitive learning for software defect prediction. IEEE Trans. Reliab. **63**(2), 676–686 (2014)
16. Jing, X., Wu, F., Dong, X., Qi, F., Xu, B.: Heterogeneous cross-company defect prediction by unified metric representation and CCA-based transfer learning. In: Proceedings of the 2015 10th Joint Meeting on Foundations of Software Engineering, pp. 496–507. ACM (2015)
17. Turhan, B., Menzies, T., Bener, A.B., Di Stefano, J.: On the relative value of cross-company and within-company data for defect prediction. Empir. Softw. Eng. **14**(5), 540–578 (2009)
18. Ma, Y., Luo, G., Zeng, X., Chen, A.: Transfer learning for cross-company software defect prediction. Inf. Softw. Technol. **54**(3), 248–256 (2012)
19. Turhan, B., Misirli, A.T., Bener, A.: Empirical evaluation of the effects of mixed project data on learning defect predictors. Inf. Softw. Technol. **55**(6), 1101–1118 (2013)
20. Peters, F., Menzies, T., Gong, L., Zhang, H.: Balancing privacy and utility in cross-company defect prediction. IEEE Trans. Softw. Eng. **39**(8), 1054–1068 (2013)
21. Peters, F., Menzies, T., Layman, L.: Lace2: better privacy-preserving data sharing for cross project defect prediction. In: ICSE'15, pp. 801–811 (2015)
22. Ryu, D., Choi, O., Baik, J.: Value-cognitive boosting with a support vector machine for cross-project defect prediction. Empir. Softw. Eng. **21**(1), 43–71 (2016)
23. Peters, F., Menzies, T., Marcus, A.: Better cross company defect predictioon. In: Proceedings of the 10th Working Conference on Mining Software Repositories, pp. 409–418. IEEE Press (2013)
24. Nam, J., Fu, W., Kim, S., Menzies, T., Tan, L.: Heterogeneous defect prediction. IEEE Trans. Softw. Eng. **44**(9), 874–896 (2018)
25. Yu, Q., Jiang, S. and Zhang, Y.: A feature matching and transfer approach for cross-company defect prediction. J. Syst. Softw. **132**, 366–378 (2017)
26. Cheng, M., Wu, G., Jiang, M., Wan, H., You, G., Yuan, M.: Heterogeneous defect prediction via exploiting correlation subspace. In: SEKE, pp. 171–176 (2016)

27. He, Z., Shu, F., Yang, Y., Li, M., Wang, Q.: An investigation on the feasibility of cross-project defect prediction. Autom. Softw. Eng. **19**(2), 167–199 (2012)
28. Ali, A., Shamsuddin, S.M., Ralescu, A.L.: Classification with class imbalance problem: a review. Int. J. Adv. Soft Comput. Appl. **7**(3), 176–204 (2015)
29. Braytee, A., Liu, W., Kennedy, P.: A cost-sensitive learning strategy for feature extraction from imbalanced data. In: International Conference on Neural Information Processing, pp. 78–86. Springer, Cham (2016)

Learning Neural Circuit by AC Operation and Frequency Signal Output

Masashi Kawaguchi, Naohiro Ishii and Masayoshi Umeno

Abstract In the machine learning field, many application models such as pattern recognition or event prediction have been proposed. Neural Network is a typically basic method of machine learning. In this study, we used analog electronic circuits using alternative current to realize the neural network learning model. These circuits are composed by a rectifier circuit, Voltage-Frequency converter, amplifier, subtract circuit, additional circuit and inverter. The connecting weights describe the frequency converted to direct current from alternating current by a rectifier circuit. This model's architecture is on the analog elements. The learning time and working time are very short because this system is not depending on clock frequency. Moreover, we suggest the realization of the deep learning model regarding the proposed analog hardware neural circuit.

Keywords Analog electronic circuit · Neural network · AC operation · Deep learning

1 Introduction

In recent years, multi-layered network models, especially the deep learning model, have been researched very actively. The performance is innovatively improved in the field of pattern/speech recognition. The recognition mechanism is elucidated more

M. Kawaguchi (✉)
Department of Electrical and Electronic Engineering, Suzuka National College of Technology, Shiroko, Suzuka Mie, Japan
e-mail: masashi@elec.suzuka-ct.ac.jp

N. Ishii
Department of Information Science, Aichi Institute of Technology, Yachigusa, Yagusa-Cho, Toyota, Japan
e-mail: ishii@aitech.ac.jp

M. Umeno
Department of Electronic Engineering, Chubu University, 1200 Matsumoto-Cho, Kasugai, Aichi 487-8501, Japan
e-mail: umeno@solan.chubu.ac.jp

R. Lee (ed.), *Computer and Information Science*, Studies in Computational Intelligence 849, https://doi.org/10.1007/978-3-030-25213-7_2

and more; self-learning IC chips have also been developed. However, these models are working on a general Von Neumann type computer with the software system.

There are only a few studies that have examined the construction of an analog parallel hardware system using a biomedical mechanism. In this research, we propose the neural network, machine learning model with a pure analog network electronic circuit. This model will develop a new signal device with the analog neural electronic circuit. In the field of neural networks, many practical use models, such as pattern recognition or event prediction, have been proposed. And many hardware implementation models, such as vision sensor or parallel network calculator, have been developed.

1.1 Analog Hardware Neural Network

The main merit of an analog hardware neural network is it can operate a continuous time system, not depending on clock frequency. On the other hand, a digital system operates depending on clock behavior on the basement of a Von Neumann type computer. As a result, advanced analog hardware models were proposed [1, 2]. In the pure analog circuit, one of the tasks is the realization of analog memory, keeping the analog value for a while [3]. The DRAM method memorizes in the capacitor as temporary memory, because it can be achieved in the general-purpose CMOS process [4]. However, when the data value is kept for the long term, it needs the mechanism to maintain the memory data. For example, a refresh process is needed. Capacitor reduces the electric charge with the passage of time. It is easy to recover the electric charge of the capacitor using a refresh process in the digital memory. However, in the case of using analog memory, the analog refresh process is quite difficult compared to the digital refresh process of DRAM. Other memorization methods are the floatage gate type device and magnetic substance memories [5, 6].

1.2 Pulsed Neural Network

Pulsed neural network receives time series pulses as learning data and changes the connection weights depending on the number of pulses. This network can keep the connecting weights of the network after the learning process and the outputs of the signal depends on the input value [7]. This network can change the connecting weights by a pulsed learning signal. However, it needs a long time for learning because many pulses are needed. For example, about no less than 1 mS is needed in the case where the pulse interval is 10 μS and 100 pulses are received to the network on the learning process.

1.3 The History of Analog Neural Network

Many research results have been reported in the field of neural network and soft computing. These research fields have been widely developed in not only pattern recognition but also control or forecasting systems. The network structure and learning method of a neural network is similar to the biomedical mechanism. The structure of the basic neural network usually consists of three layers. Each layer is composed of the input, connecting weight, summation, thresholds and output unit. In our previous study, we designed and proposed motion detection system or image processing model using a multi-layered neural network and an artificial retina model.

On the other hand, we construct a pattern recognition machine using variable resistance, operational amplifier. We used CdS cells in the input sensor. However, we have to adjust the resistance value by our hands. Moreover, the capacitor reduces the electric charge with time passage. It needs the analog refresh process. The analog refresh process is quite difficult compared to the digital refresh process of DRAM. In the present study, we proposed a neural network using analog multiple circuits and an operational amplifier. The learning time and working time are very short because this system is not dependent on clock frequency in the digital clock processing. At first we designed a neural network circuit by SPICE simulation. Next we measured the behavior of the output value of the basic neural network by capture CAD and SPICE. Capture is one of the computer-aided design (CAD) systems of SPICE simulation. We compared both output results and confirmed some extent of EX-OR behavior [8, 9]. EX-OR behavior is typically a working confirmation method of a three layered neural network model. This EX-OR behavior makes liner separation impossible, it is suitable data for checking neural network ability.

Moreover, the model which used the capacitor as the connecting weights was proposed. However, it is difficult to adjust the connecting weights. In the present study, we proposed a neural network using analog multiple circuits. The connecting weights are shown as a voltage of multiple circuits. It can change the connecting weights easily. The learning process will be quicker. At first we made a neural network by computer program and neural circuit by SPICE simulation. SPICE means the Electric circuit simulator as shown in the next chapter. Next we measured the behavior confirmation of the computer calculation and SPICE simulation. We compared both output results and confirmed some extent of EX-OR behavior [8].

2 Neural Network Using Multiple Circuit

In our previous study, we used multiple circuits to realize the analog neural network. In the SPICE simulation, the circuit is drawn by CAD, called Capture. After setting the input voltage or frequency, SPICE has some analysis function, AC, DC or transient. At first, we made the differential amplifier circuits and Gilbert multiple circuits toward making the analog neural network. We show the different circuits in

Fig. 1 Difference circuits

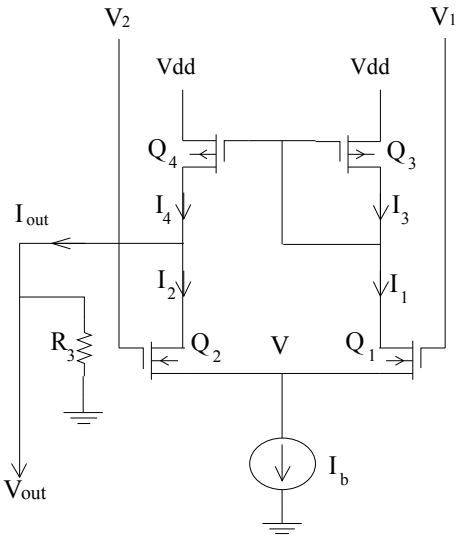


Fig. 1. Many circuits take an input signal represented as a difference between two voltages. There circuits all use some variant of the differential pair. Figure 1 shows the schematic diagram of the simple transconductance amplifier with differential pairs. The current mirror formed by Q3 and Q4 is used to form the output current, which is equal to I_1–I_2. The different circuits enhanced to a two-quadrant multiplier. Its output current can be either positive or negative, but the bias current I_b can only be a positive current. V_b, which controls the current, can only be positive voltage. So the circuit multiplies the positive part of the current I_b by the tanh of (V_1–V_2). If we plot V_1–V_2 horizontally, and I vertically, this circuit can work in only the first and second quadrants.

We show the Gilbert multiple circuit in Fig. 2. To multiply a signal of either sign by another signal of either sign, we need a four-quadrant multiplier. We can achieve all four quadrants of multiplication by using each of the output currents from the differential pair (I_1 or I_2) as the source for another differential pair. Figure 2 shows the schematic of the Gilbert multiplier. In the range where the tanh x is approximately equal to x, this circuits multiplies V_1–V_2 by V_3–V_4. And we confirmed the range of the voltage operated excellently. One neuron is composed by connecting weights, summation and threshold function. The product of input signal and connecting weights is realized by multiple circuits. Summation and threshold function are realized by additional and difference circuits. In the previous hardware model of neural network, when we use solid resistance elements, it needs to change the resistance elements with the complex wires in each step of learning process. In the case of using variable resistance, we have to adjust the resistance value by our hands.

Figure 3 is the one neuron circuit, using multiple circuits and an additional circuit by opamp. Multiple circuits calculate the product of the two input values, input signals and connecting weights. There are three multiple circuits. Two multiple circuits mean

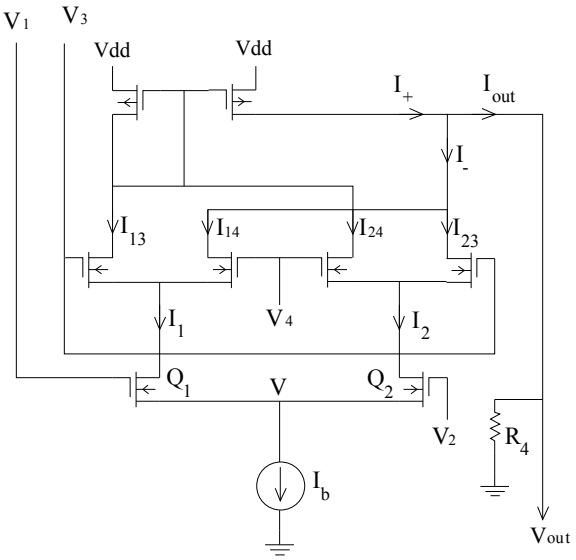


Fig. 2 Gilbert multiple circuits

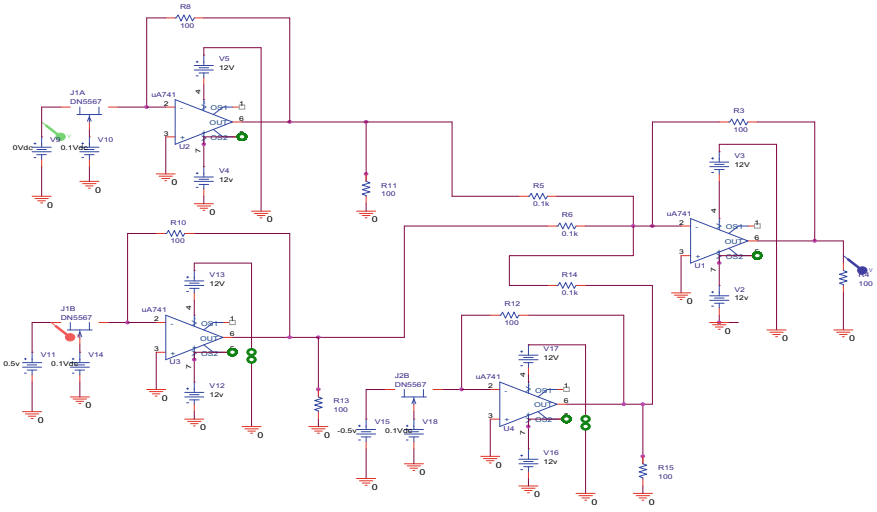


Fig. 3 Neural circuit by capture CAD

two input signals and connecting weights. The other one multiple circuit means the threshold part of basic neuron. In the threshold part, the input signal is −1. In the multiple circuit, its products input signal −1 and connecting weights. So the output of the multiple circuit is the threshold of this neuron.

3 Perceptron Feedback Network by Analog Circuits

In Fig. 4, we show the architecture of the perceptron neural network. This is a basic learning network using a teaching signal. 'y' means the output signal of the neural network. 't' means the teaching signal. Error value 't − y' is calculated by the subtract circuit. After calculating the error value, the product error value and input signal are calculated and make a feedback signal. The input of the subtract circuit on the feedback line are the feedback signal and the original connecting weight. This subtract circuit calculates the new connecting weight. After the product of the new connecting weigh is calculated, the next time learning process is started.

Figure 5 shows the perceptron circuits, two-input and one-output. There are multiple circuits and additional circuits in the feed forward line. Error value between original output and teaching signal is calculated by subtract circuit. There are multiple circuits and additional circuit in the feedback lines. In the experimental result of this perceptron, the learning time is about 900 μs shown in Fig. 6 [10]. Figure 7 shows the Architecture of Three-Layers Neural Circuits. In Fig. 8, we show the Learning Neural Circuit on Capture CAD by SPICE.

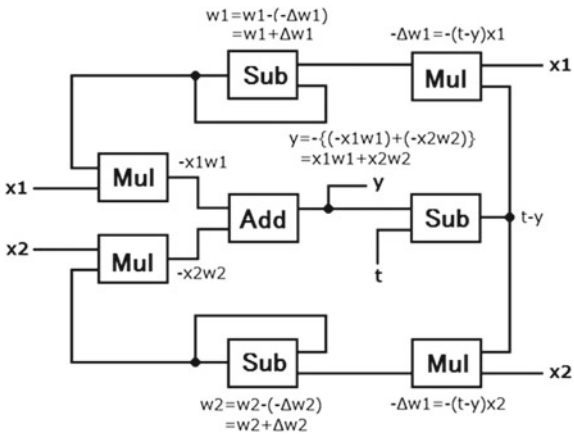


Fig. 4 The architecture of perceptron

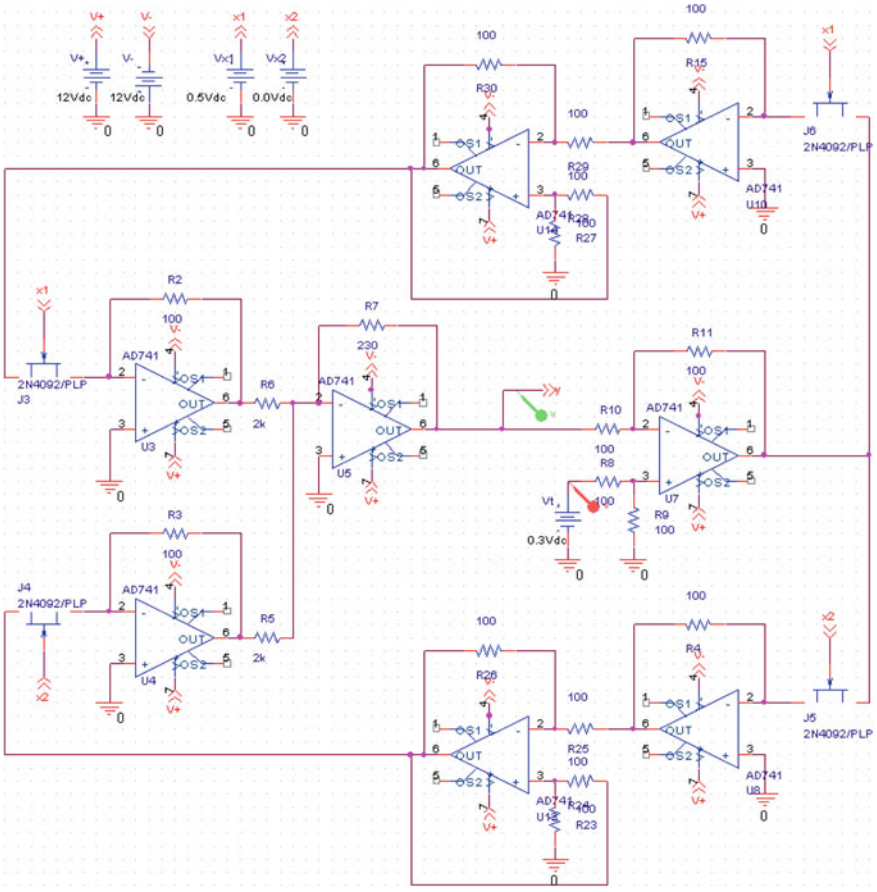


Fig. 5 The circuit of perceptron

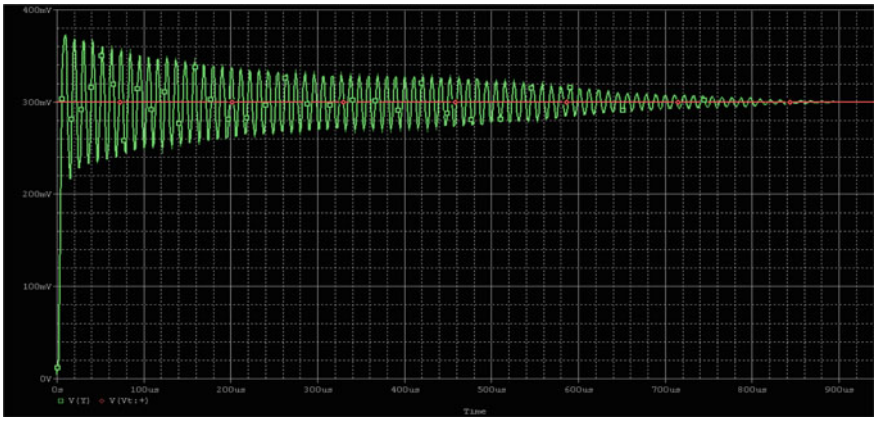


Fig. 6 The convergence output of perceptron

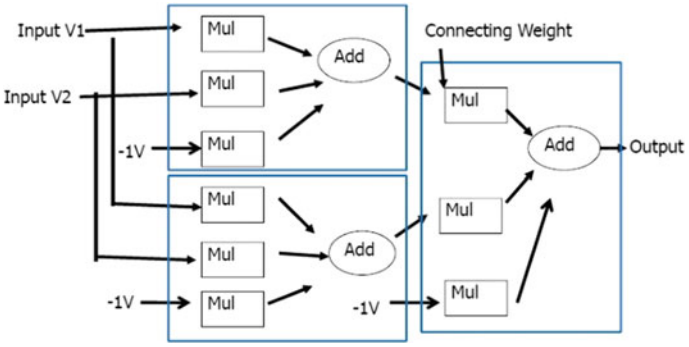


Fig. 7 The diagram of neural circuits with threshold

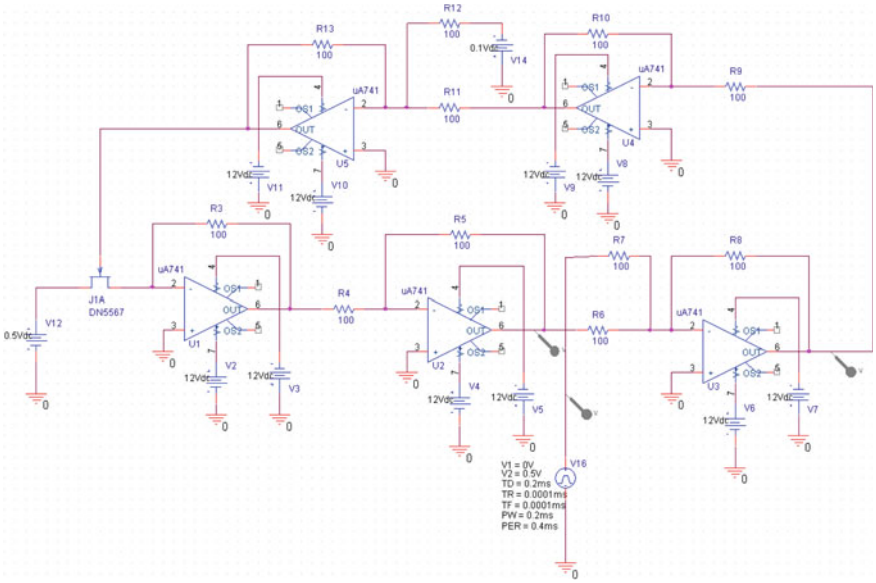


Fig. 8 The learning feedback neural circuit

4 Neural Circuit on Alternative Current Behavior

We proposed an analog neural network using multiple circuit in the previous research. However, in the case of constructing a network, one of the disadvantages is that the input and output range is limited. Furthermore, the circuit operation becomes unstable because of the characteristics of the multiple circuit using semiconductor. It is called 'Circuit Limitations'. One of the cause is transistor mismatch. Not all transistors created equal. Another cause is the output-voltage limitation. We tried to use the alternative current as a transmission signal in the analog neural network in Fig. 9.

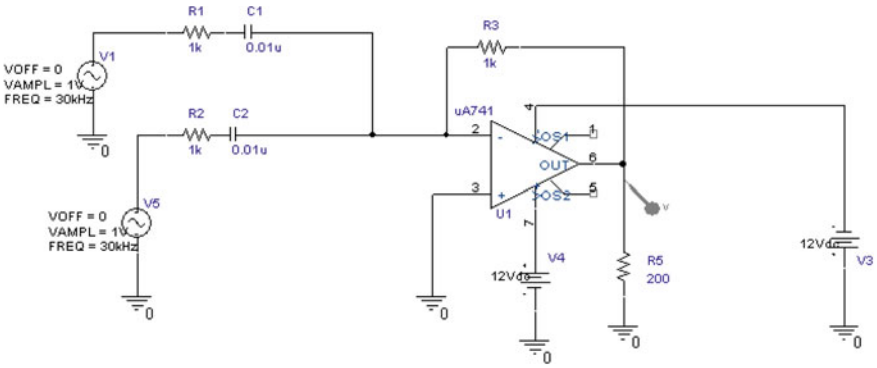


Fig. 9 AC operation neural circuit

The original input signal is direct current. We used the voltage frequency converter unit when generate the connecting weight. The input signal and connecting weight generate the Alternative current by the Amplifier circuit. Two Alternative currents are added by an additional circuit. The output of the additional circuit is a modulated wave. This modulated wave is the first output signal of this neural network.

Figure 10 is the output of the RMS value of AC voltage by the neural circuit. In this network, two Alternative currents are added by an additional circuit. The output of the additional circuit is a modulated wave. Figure 10 shows the RMS value of

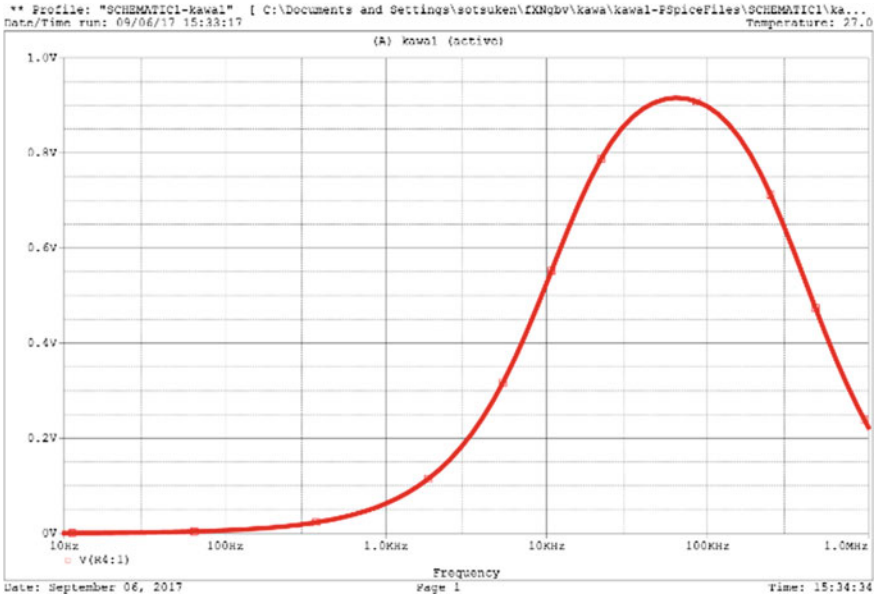


Fig. 10 The output Rms Value of neural circuit

the modulated wave. It operates satisfactorily because the output voltage increases monotonically in the general-purpose frequency range. Figure 11 is the output of the neural circuit. It is shown by two dimensional graph. We recognized the RMS value of the output voltage is the appropriate value in the two-dimensional area.

When we construct learning the AC neural circuit, we have to convert the feedback modulated current signal to a connecting weight with frequency. The correction error signal is calculated by the products difference signal and input signal. Difference signal is the difference between the output value and teaching signal. Figure 12 shows the convergence result of learning experiment. It means the learning process is succeeding with very short time. Figure 13 shows the Basic AC operation learning neural network model. This circuit is composed by a Rectifier circuit, Voltage-Frequency converter, Amplifier, subtract circuit, Additional Circuit and Inverter. The input signal is direct current. The initial value of the connecting weight is also direct current. This direct current is converted to frequency by a Voltage-Frequency converter circuit. The input signal and connecting weight generate the Alternative current by the Amplifier circuit (Fig. 14).

Figure 15 shows the relationship between the number of learning time and output frequency. It shows the frequency f1 convergences to 4 kHz and the frequency f2

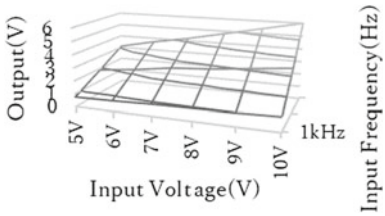


Fig. 11 The output behavior of AC operation neural circuit

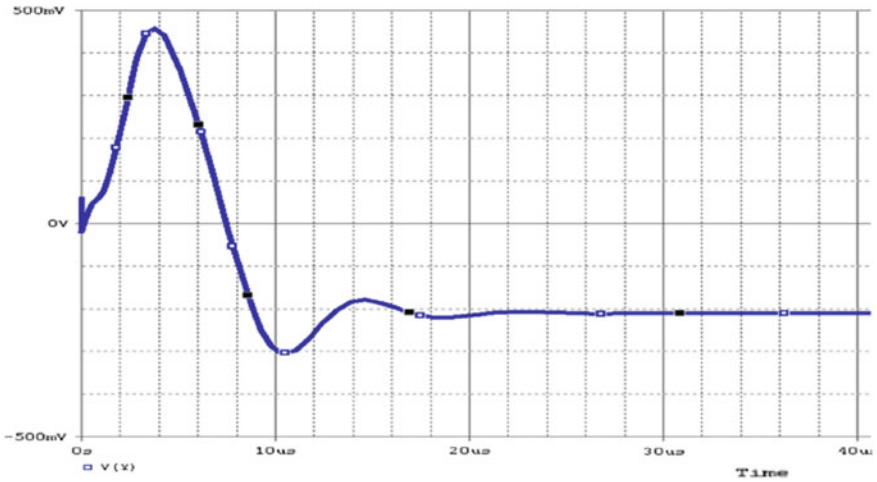


Fig. 12 The convergence result of learning experiment

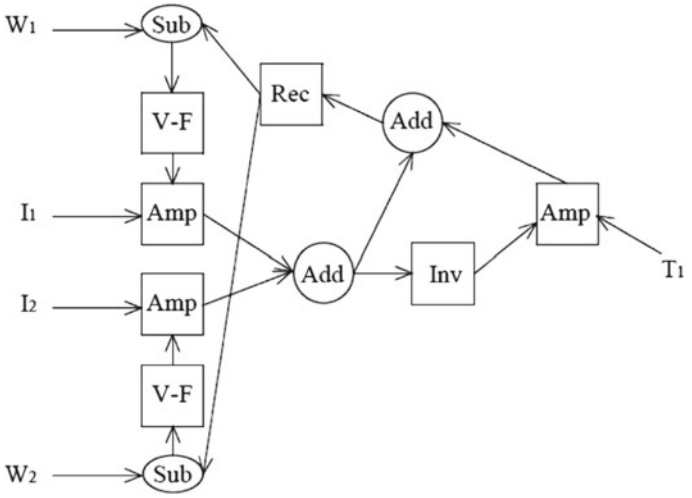


Fig. 13 Basic AC operation learning neural network model

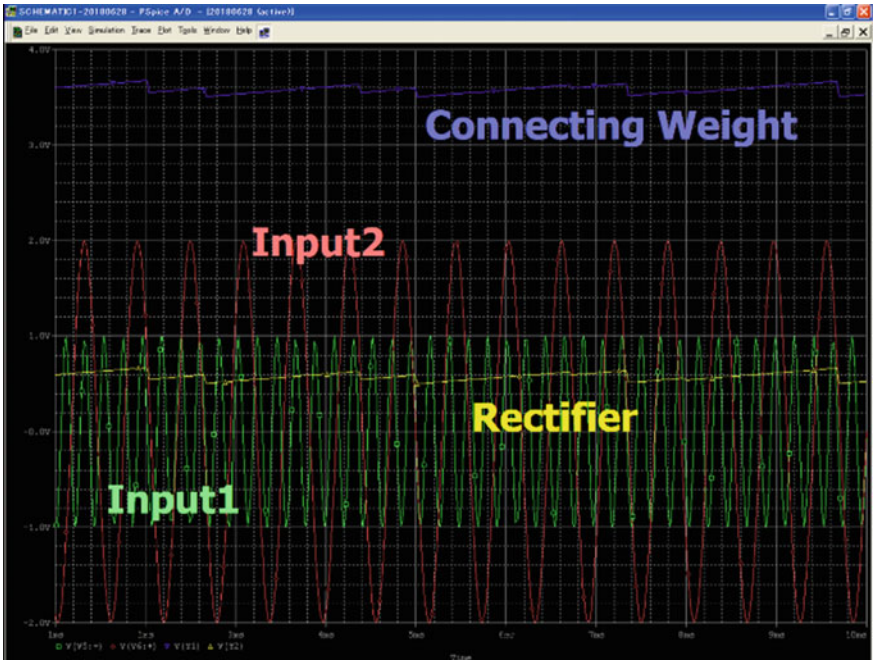


Fig. 14 The simulation results of AC feed-back neural model

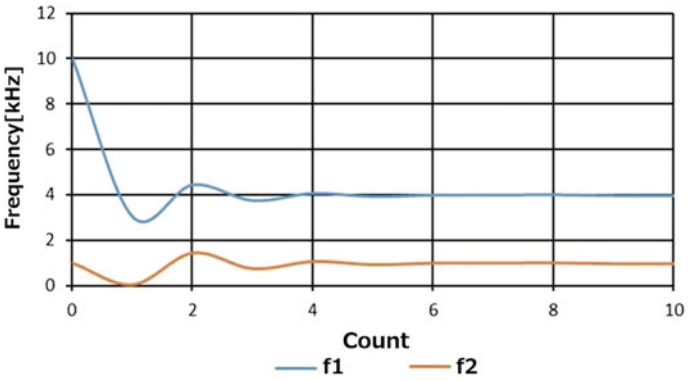


Fig. 15 The number of learning time and output frequency

convergences to 1 kHz. The learning count time is very small. The learning speed is very fast in this AC operation circuit. Figure 16 shows the whole circuit of AC operation learning neural network.

Two alternative currents are added by an additional circuit. The output of the additional circuit is a modulated wave. This modulated wave is phase inverted by

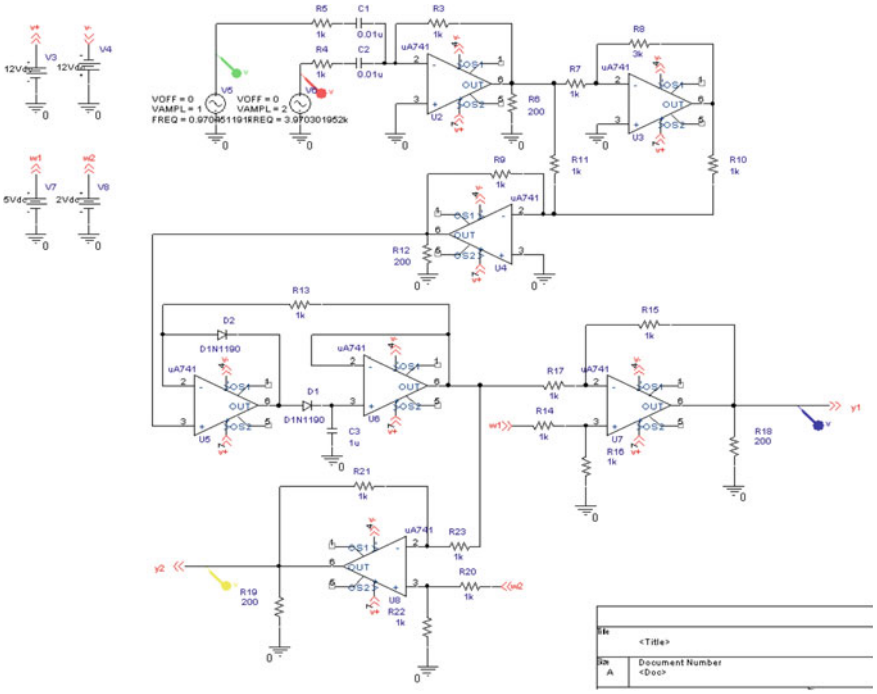


Fig. 16 The AC operation learning neural circuit

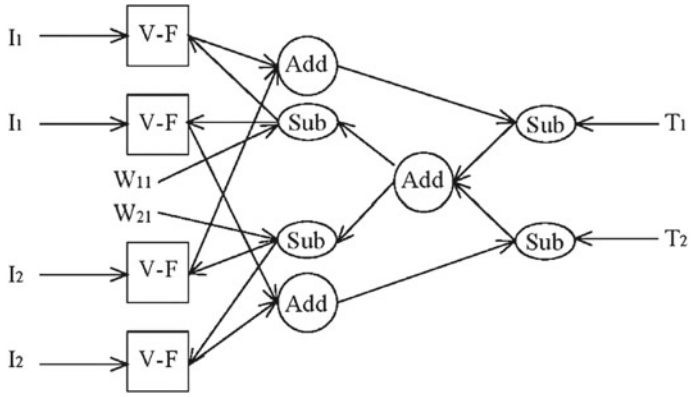


Fig. 17 The structure of 2-patterns analog AC operation neural network with V-F conversion circuit

an Inverse circuit. The phase-inverted wave is amplified. The amplification is the value of the teaching signal. This amplified signal and modulated wave are added by an adder circuit. The output of this adder circuit is the error value, which is the difference between the output and teacher signal. Thus, we do not have to use the subtract circuit to calculate the error value (Figs. 17 and 18).

The output of the Adder circuit is converted to direct current from alternating current by the rectifier circuit. This direct current is a correction signal of connecting weights. New connecting weight is calculated by a subtract circuit. This circuit calculates the original connecting weight and the correction signal of the connecting weight. The output of the subtract circuit is converted to a frequency signal by a voltage-frequency convert circuit. It means that in the AC feedback Circuit for the BP learning process, after getting DC current by a rectifier circuit, we have to convert from DC Voltage to AC current with frequency. Finally, alternating current occurs by the amplifier circuit. The amplification is the value of the input signal. Figure 14 shows the simulation results of AC feed-back neural model, two-input signal, connecting weights and after rectified wave.

5 Deep Learning Model

Recently, a deep learning model has been proposed and developed in many applications such as image recognition and artificial intelligence. Deep learning is a kind of algorithms in the machine learning model. This model is developed in the recent research. The recognition ability is improved more and more. Not only pattern recognition, but also in image or speech recognition fields, the deep learning model is used in many fields in practical use. And this system is expected in the field of robotics, conversation system and artificial intelligence.

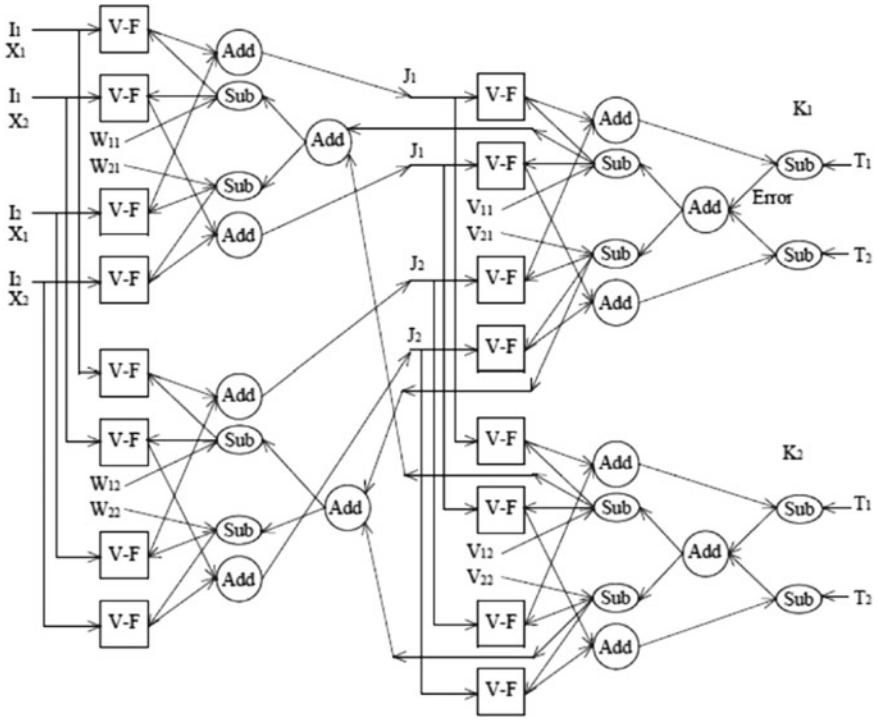


Fig. 18 The structure of enhanced analog neural network to three layers of Fig. 15

5.1 The Stacked Auto Encoder

The stacked auto-encoder is one of the processes in deep learning. This is the pre-learning method of the large number layer network. In the basic neural network, in the almost case, there are three layers. On the other hand, there are nine layers in the deep learning model generally. After the learning process is completed by machine learning method, remove the decoding part, output layer and the connection of intermediate layer and output layer. The keeping of the coded portion means from the input layer to the intermediate layer including the connection of the input layer and intermediate layer. The intermediate layer contains the compressed data of the input data. Moreover, we obtain more compressed internal representation, as the compressed representation input signal to apply the auto-encoder learning. After removing the decoding part of stacked auto-encoder, next network is connected. This network is also learned by another three-layered network and remove the decoding part. A Stacked auto-encoder has been applied to the Restricted Boltzmann Machine (RBM) as well as the Deep learning network (DNN). Moreover, stacked auto-encoder is used the many types of learning algorithm. Recently, the learning experiment featuring a large amount of extraction from an image has become well-known. Stacked

auto-encoder can self-learning of abstract expression data. This network has nine layers with three superimposed sub-networks, such as a convolution network [11]. In the previous research, we described the simple neural network learning model by analog electronic circuits. We tried to expand the network to realize the deep learning model. Next, we constructed 2 input, 1 output and 2 patterns neural model as in Fig. 15. In this circuit, each pattern needs each circuit. For example, in the case there are 5 kinds of learning patterns, we have to construct 5 input unit circuits. However, learning time is very short. "V-F" means the V-F converter circuit. The output of the subtract circuit is converted to a frequency signal by a voltage-frequency convert circuit in Fig. 17.

The output of the subtract circuit is converted to a frequency signal by a voltage-frequency convert circuit. It means that in the AC feedback Circuit for BP learning process, after getting DC current by rectifier circuit, we have to Convert from DC Voltage to AC current with Frequency. I_1 and I_2 are input units. Two I_1 mean two inputs. T_1 and T_2 means two teaching signals. W_{11} and W_{12} are connecting weights. Figure 18 means the expand network of Fig. 17. Although this model needs many neural connections, the learning speed is very high because of the plural data patterns learning occurs at the same time and working analog real time system not depending on clock frequency. And after learning, each new connecting weight between the input layer and middle layer is picked up, it is the parted potion including the connecting weights between the input layer and middle layer as well as the layers of input and middle. It means the stacked auto encoder process and suggests the possibility of design of many layers of the deep learning model [12]. To fix the connecting weights after learning process, we proposed the two-stage learning process. In the learning stage, connecting weighs are able to change depending on the teaching signal. After learning process finished, we used the sample hold circuit to fix the connecting weights. In this situation, this circuits receive the input signal and outputs the output signal in the environment that all the connecting weights are fixed.

6 Conclusion

At first, we designed an analog neural circuit using multiple circuits. We confirmed the operation of this network by SPICE simulation. Next, we constructed a basic analog neural network by an alternative current operation circuit. The input signal and connecting weight generate the alternative current by the amplifier circuit. Two alternative currents are added by an additional circuit [13, 14]. The frequency signal is generated by a Voltage-Frequency converter circuit. The input signal of the V-F converter is rectified direct current. The input of the rectified circuit is the error correction signal by alternative current. The connecting weight can be changed by an error-correction signal and the input frequency is depending on the output Voltage-Frequency converter circuit in the feedback learning process. This model has extremely high flexibility characteristics. It is the AC feedback circuit for the BP learning process, after getting DC current by rectifier circuit, we have to convert from

DC Voltage to AC current with frequency. Moreover, a deep learning model has been proposed recently and developed in many applications such as image recognition and artificial intelligence. In the future, this hardware learning system is expected in the field of robotics, conversation systems and the artificial intelligence.

References

1. Mead, C.: Analog VLSI and Neural Systems. Addison Wesley Publishing Company, Inc. (1989)
2. Chong, C.P., Salama, C.A.T., Smith, K.C.: Image-motion detection using analog VLSI. IEEE J. Solid-State Circuits **27**(1), 93–96 (1992)
3. Lu, Z., Shi, B.E.: Subpixel Resolution Binocular Visual Tracking Using Analog VLSI Vision Sensors. IEEE Trans. Circuits Syst.-II Analog Digit. Signal Process. **47**(12), 1468–1475 (2000)
4. Saito, T., Inamura, H.: Analysis of a simple A/D converter with a trapping window. In: IEEE International Symposium on Circuits and Systems, pp. 1293–1305 (2003)
5. Luthon, F., Dragomirescu, D.: A cellular analog network for MRF-based video motion detection. IEEE Trans Circuits Syst.-I Fundam. Theory Appl. **46**(2), 281–293 (1999)
6. Yamada, H., Miyashita, T., Ohtani, M., Yonezu, H.: An analog MOS circuit inspired by an inner retina for producing signals of moving edges. Technical Report of IEICE, NC**99**-112, 149–155 (2000)
7. Okuda, T., Doki, S., Ishida, M.: Realization of back propagation learning for pulsed neural networks based on delta-sigma modulation and its hardware implementation. ICICE Transactions, J**88**-D-II-4, 778–788 (2005)
8. Kawaguchi, M., Jimbo, T., Umeno, M.: Analog VLSI layout design and the circuit board manufacturing of advanced image processing for artificial vision model. In: KES2008, Part II, LNAI, vol. 5178, pp. 895–902 (2008)
9. Kawaguchi, M., Jimbo, T., Umeno, M.: Dynamic Learning of Neural Network by Analog Electronic Circuits. In: Intelligent System Symposium, FAN2010, S3-4-3 (2010)
10. Kawaguchi, M., Jimbo T., Ishii, N.: Analog learning neural network using multiple and sample hold circuits. In: IIAI/ACIS International Symposiums on Innovative E-Service and Information Systems, IEIS 2012, pp. 243–246 (2012)
11. Yoshua, B., Aaron, C., Courville, P.: Vincent: representation learning: a review and new perspectives. IEEE Trans. Pattern Anal. Mach. Intell. **35**(8), 1798–1828 (2013)
12. Kawaguchi, M., Ishii, N., Umeno, M.: Analog neural circuit with switched capacitor and design of deep learning model. In: 3rd International Conference on Applied Computing and Information Technology and 2nd International Conference on Computational Science and Intelligence, ACIT-CSI, pp. 322–327 (2015)
13. Kawaguchi, M., Ishii, N., Umeno, M.: Analog learning neural circuit with switched capacitor and the design of deep learning model. In: Computational Science/Intelligence and Applied Informatics, Studies in Computational Intelligence, vol. 726, pp. 93–107 (2017)
14. Kawaguchi, M., Ishii, N., Umeno, M.: Analog neural circuit by AC operation and the design of deep learning model, DEStech transactions on computer science and engineering. In: 3rd International Conference on Artificial Intelligence and Industrial Engineering, pp. 228–233 (2017)
15. Kawaguchi, M., Jimbo, T., Umeno, M.: Motion detecting artificial retina model by two-dimensional multi-layered analog electronic circuits. IEICE Trans. E**86**-A-2, 387–395 (2003)
16. Kawaguchi, M., Jimbo, T., Umeno, M.: Analog VLSI layout design of advanced image processing for artificial vision model. In: ISIE2005 Proceedings of the IEEE International Symposium on Industrial Electronics, vol. 3, pp. 1239–1244 (2005)

Preprocessing with Contrast Enhancement Methods in Bone Age Assessment

Aye Aye Aung and Zin Mar Win

Abstract Bone age assessment (BAA) using radiological x-rays of the left-hand wrist is important in pediatric endocrinology to correctly assess growth and pubertal maturation. To identify the age, the most commonly used technique is bone age assessment. The bone age is the most commonly used criteria in age and growth disorder. There are two main methods for bone age estimation such as Tanner & Whitehouse (TW) method and Greulich & Pyle (GP) method. Tanner & Whitehouse (TW) method is a score assigning method while Greulich & Pyle (GP) is an atlas matching method. The objective of this research is to improve the accuracy of the automated bone age estimation. To improve the robustness and accuracy of this system, image preprocessing techniques are also necessary. Bone age assessment can affect the accuracy in segmentation because of poor contrast, noise, and various hand positions. Preprocessing steps of this system include background removal, radiological marker removal, image enhancement, and hand rotation methods. In this paper, we use the Contrast Limited Adaptive Histogram Equalization (CLAHE) method to enhance the contrast of the hand bone radiograph. This method enhances the hand radiograph in which the background and body have more brightness. The advantage of CLAHE is to prevent the over-amplification of noise, it is able to increase contrast. In this paper, we tested contrast enhancement techniques such as Contrast Limited Adaptive Histogram Equalization (CLAHE), Histogram Equalization (HE) and Power Law Transform (PLT) techniques.

Aye Aye Aung (✉) · Zin Mar Win
Image Processing Lab, University of Computer Studies, Mandalay, Myanmar
e-mail: ayeayeaungucsm86@gmail.com

Zin Mar Win
e-mail: zinmarwin@ucsm.edu.mm

R. Lee (ed.), *Computer and Information Science*, Studies in Computational Intelligence 849, https://doi.org/10.1007/978-3-030-25213-7_3

1 Introduction

The input images have to be enhanced to analyze the process in the system. Raw images may have noise and so it is required to enhance. Enhancement of the image quality is obtained by implementing filtering technique, removal of noise and contrast enhancement methods [1].

Bone age assessment is a common clinical method to diagnose metabolic disorders and endocrine in children.[1] Bone age from left-hand wrist radiograph is assessed and then compared with chronological age which is calculated from the date of birth [2]. Based on the difference between the bone age and chronological age, pediatricians can diagnose the child' health problems such as growth disorder, endocrine diseases, and diabetes. The bone age cannot be computed from hand bone radiograph at the age of 18 years old because the hand is matured at 18 years. Therefore, the clavicle is used to estimate the age in age 18–22-year-olds. The discrepancy between chronological and bone age indicates abnormalities. Height, bone maturity and dental age describe Human's chronological situations. Bone age estimation is important because it is reliable and practicable in growth disorders and endocrinology. Race, nutrition, environmental status, gender influence the bone development of children.

Tanner & Whitehouse (TW) method and Greulich & Pyle (GP) method are two main clinical methods in automated bone age estimation system. GP method is faster and easier than the TW method. TW method is a more accurate estimation than the GP method and more suitable for automation. A left-hand bone image is compared with radiograph in atlas according to sex and age in the GP method. Twenty regions of interest (ROIs) of hand radiograph is considered to assess bone age in TW method. Each ROI consists of three parts such as Epiphysis, Metaphysis, and Diaphysis. The development stages of each ROI are divided into A, B, C, . . ., I stages. Each stage of each bone is associated with a numerical score of bone. The total maturity score is obtained by summing all ROI scores. These scores have a direct correlation with the bone age separately depending on gender and race. Hand wrist bone radiograph is the commonest method for child bone age assessment. A left-hand radiograph is suitable to assess bone age because hand possesses many bones and taking a radiograph of hand is easy.

The proposed system has the following steps such as preprocessing step, segmentation step, feature extraction step, and classification step. In this paper, we implement preprocessing steps of bone age estimation system. In the preprocessing step, contrast enhancement method, background removal and hand rotation method are also included. As a contrast enhancement method, we apply the CLAHE method to enhance the contrast of hand radiograph. In background removal, Otsu's thresholding method is applied. Finally, hand rotation method is implemented.

[1] https://en.wikipedia.org/wiki/Adaptive_histogram_equalization.

2 Literature Review

The aim of the research is to assess automated skeletal bone age estimation with Epiphysis/Metaphysis ROI of the middle finger and carpal ROIs. Therefore, we study existing skeletal bone age estimation techniques. Bone age assessment is frequently used in pediatric endocrinology to evaluate growth and diagnose growth disorder. With the development of image processing technique, many investigators have attempted bone age assessment with the help of image processing. In the 19th century, dentists estimated age and then to estimate a child's age, tooth eruption was considered as a reliable method. In that century, the bone age was calculated from seven years old as a minimum in Britain [3]. Some researchers have opposed this method to estimate age. Dr. Pedro Mata announced about age estimation concern with tooth development in 1846 [4]. In 1895, Rontgen' discovery made a revolution to estimate an age for living subjects. The innovation of skeletal radiograph was used as a tooth eruption' complement [5]. In 1886, Angerer found that the carpal bones of the wrist are an indicator to estimate the age of children.

Zhang et al. [6] developed an automated bone age estimation system. They used fuzzy classification for bone age estimation according to the features extracted from carpal bones. They used 205 hand wrist images by two radiologists' readings. It was found that the carpal bones are very reliable to determine the bone age from 0 to 7 years old children. Giordano et al. [7] proposed an automated bone age estimation by using Epiphysis/Metaphysis (EMROI) and carpal ROI (CROI) analysis together based on Tanner and Whitehouse (TW2) method. They evaluated bone age estimation in the age range from newborn to 10 years old of children. In this paper, bone extraction is carried for carpal ROI analysis whereas using active contour models and derivative difference of Gaussian filter. They proposed different of Gaussian filter for EMROI analysis. In this paper, the experimental result was achieved by about 87% on 30 hand images. The performance of this system was degraded about 50% by using Trapezium and Trapezoid. Without considering these bones the accuracy of this paper raised to 91%.

In 2012, Thangam et al. [8] proposed automated bone age estimation from epiphysis/metaphysis region of interest (EMROI) by using Hausdorff distance classification. This system was based on TW2 method of bone age assessment. This method was evaluated bone age from age 0 to 10-year-olds for both male and female. Guraksin et al. [9] proposed an automated bone age assessment system using on carpal and radius bone. They used the C4.5 algorithms, support vector machine, naive Bayes and k-nearest neighborhood for classification. They assessed bone age estimation in the age range from newborn to 6 years in applying 195 X-ray images. They used nine features of carpal and distal epiphysis of the radius bone. The best six features form gain ratio was applied for bone age classification. The performance of this system showed that the support vector machine method has a better accuracy rate than others at 72.82%. Choi et al. [10] proposed a simple bone age assessment method called capitohamate (CH) planimetry. CH planimetry was measured by summing the areas of capitate and hamate of carpal bones. CH planimetry has used both hands

of children. In this paper, CH planimetry was compared with Greulich & Pyle (GP) method. They assessed bone age estimation in the age range from 0 to 180 months in 109 children.

Gertych et al. [11] proposed an automated bone age assessment method of children based on a digital hand atlas. They used 1400 left hand radiographs from Caucasian, Asian, African-American and Hispanic children range from one to eighteen year olds. They extracted features from seven ROIs such as the six phalangeal ROIs and carpal bones. They used fuzzy classifiers to train the extracted features.

In 2008, Liu et al. [12] proposed particle swarm optimization (PSO) method for bone age segmentation and feature extraction. They used 5 features of each ROI, including morphological characteristics, size, and fusional/adjacent stage. This system applied Artificial Neural Network (ANN) classifiers by using extracted features. They applied ANNs to multilayer feed-forward networks and trained in processing extracted bone features with back-propagation algorithm rules. They trained about 1046 left hand radiographs on ANNs. According to their results, PSO method was more effective and accurate in image segmentation and feature extraction.

In 1993, Pietka et al. [13] performed automated bone age assessment based carpal bones analysis. They used standard thresholding method to separate hand from the background. Dynamic thresholding method was used to differentiate the bone from soft tissues. This system used features such as area, perimeter, compactness ratio, the center of gravity, coefficient of convexity, lengthening ratio, average gray pixel value, and discrepancy of the gray pixel value. Seok et al. [14] in 2012 proposed a fully automated bone age estimation based on Greulich and Pyle Atlas (GP) using feature extraction and classifiers for machine learning. They used the Scale Invariant Feature Transform (SIFT) to extract features from the images then apply Singular Value Decomposition (SVD) to compose the feature vectors. Then, they train a Neural Network classifier using extracted feature vectors. According to their results, even with a small number of training data sets, they obtain promising results. In 2017, Spampinato et al. [15] proposed automated bone age assessment using deep learning. They tested several existing pre-trained convolutional neural networks such as OverFeat, GoogleNet and OxfordNet on 1400 X-ray images. They also designed and trained a CNN also called BoNet (including five convolutional layers), which is the most effective and robust.

In 2016, Giordano et al. [16] proposed automated bone age assessment system to determine the bone age based on left hand wrist X-rays according to Tanner and Whitehouse method. They used Epiphysis, Metaphysis (EMROIs) only and assessed bone age estimation in age range 0–6 years. This research applied Hidden Markov Model (HMM) based classifiers by using extracted feature vectors. According to their experimental results, they achieved accuracy rate of 95% in applying 360 X-ray images.

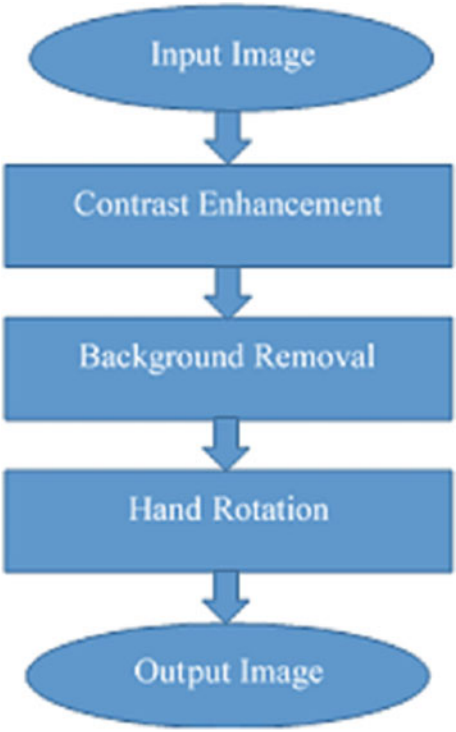


Fig. 1 Processing steps of the research

3 Material and Method

The raw images may have unwanted noises. These images are not suitable for direct processing because of unwanted noises present in these images. So, the input images are necessary to preprocess. Preprocessing is an important step used in MRI, label, artifact removal, enhancement, and segmentation [17]. The preprocessing involved in the conversion, image resize, the noise removing and enhances the quality and produces an image in which minutiae can be detected correctly [18]. Because of poor contrast, radiological markers and unwanted noises, it can affect the accuracy of segmentation negatively. In this research work, the preprocessing system includes contrast enhancement, background removal, radiological markers removal, and hand rotation. Figure 1 shows the preprocessing steps of the research work.

The system is an estimate of bone age based on the method of Tanner and Whitehouse (TW2). Therefore, we want to describe Tanner and Whitehouse method.

3.1 Tanner and Whitehouse (TW2) Method

Tanner & Whitehouse method is a score based bone age estimation method. Tanner and Whitehouse method employs twenty ROIs located on the first, middle and fifth fingers and the carpal bones (including radius and ulna) of the left-hand wrist. Each ROI has three parts such as Epiphysis, Metaphysis, and Diaphysis. In this method, the bones are divided into eight or nine discrete development stages of hand wrist. Each bone of hand has various developmental stages and each developmental stage has various criteria Each stage is assigned into a letter from A, B, C, D, E, F, G, H stages in which A means the sign of bone is absent. Each stage of each bone is associated with a maturity score. These scores of the development stages are gender specific. The sum of these scores is in the range from 0 (invisible) to 1000 (maturity). The total maturity score is related to bone age. Figure 2 shows the hand and wrist radiograph and Fig. 3 shows metaphysis, epiphysis and diaphysis ROIs of hand bone.

3.2 Contrast Enhancement

Contrast enhancement is one of the interesting technique used for enhancement of contrast of an image. Contrast enhancement enhances the brightness difference between objects and their backgrounds so it improves the perceptibility of objects [19]. Contrast enhancement techniques are important for both human and computer vision in image processing. Contrast enhancement methods adjust the brightness and darkness of objects to improve their visibility. Firstly, we apply the CLAHE method to enhance the input x-ray. CLAHE is a variation of adaptive histogram equalization that reduces the noise amplification problem by limiting the contrast amplification.[2] CLAHE method is usually applied to enhance the low contrast medical images. CLAHE technique is effective in adjusting the contrast of the hand radiograph. This method can lead to better views of bones in hand radiograph. This method enhances the hand radiograph in which the background and body have more brightness. CLAHE method divides the image into contextual regions and then applies the equalization of histograms to each region. In this paper, we implement the CLAHE method by comparing Power Law Transform (PLT) and Histogram Equalization (HE) methods.

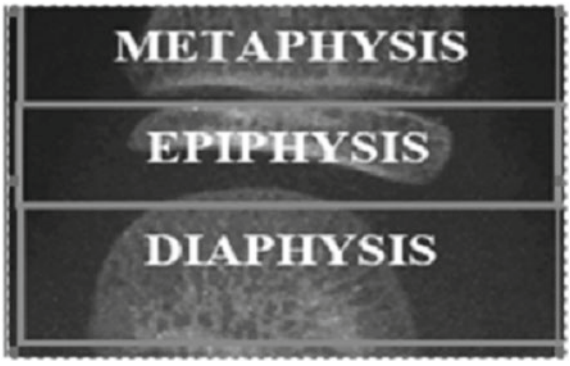
Histogram Equalization is a technique that adjusts the contrast of the image using the image' histogram. Histogram Equalization uniformly distributes grey level of the image' histogram to obtain the good contrast of an image.

Power Law Transformation merges together foreground and background peaks of low contrast image. Figure 4a, b, c, d are compared input x-ray with CLAHE, HE, and PLT methods.

[2] https://en.wikipedia.org/wiki/Adaptive_histogram_equalization.

Fig. 2 Hand and wrist bones radiograph

Fig. 3 Metaphysis, Epiphysis and Diaphysis ROI



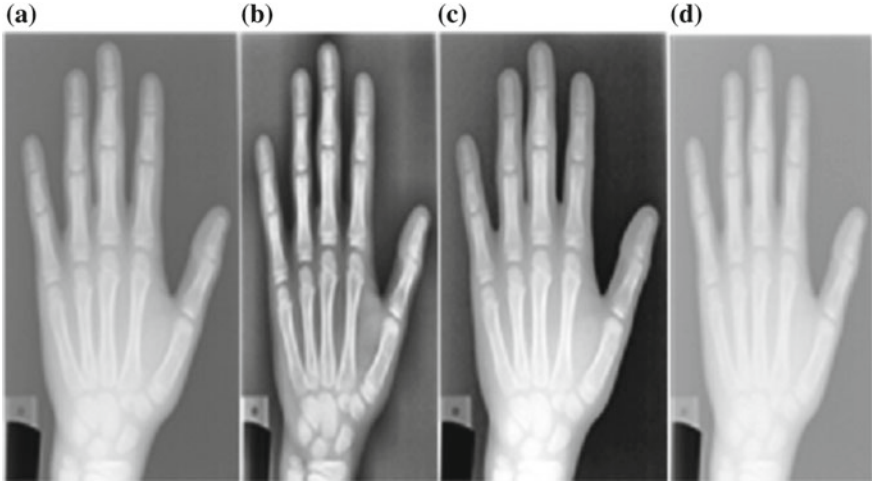


Fig. 4 **a** Input Image. **b** CLAHE Image. **c** HE Image. **d** PLT Image

3.3 Background Removal

Some images require a particular grayscale adjustment step. Ephysis/metaphysis ROI and carpal ROI analyses need to increase the ratio of hand to the background which is important when an under or overexposed image is the input of the system.[3] In background removal, the input image is thresholded using Otsu's method. It is the reduction of grayscale image to binary image. In Otsu's algorithm, the image has two main classes of pixels such as foreground and background pixels. It calculates a threshold to separate into two classes. To remove small noisy elements such as radiological markers, we use the label connected component function of the 2-D binary image. This function assigns the pixels of all smaller objects into zero and the largest object into one. Figure 5 shows the result of thresholded hand images obtained from Otsu' method.

3.4 Hand Rotation

Hand orientation is a standard image processing method to extract the region of interest (ROIs). Hand orientation is necessary because x-ray images are often not well oriented. Hand rotation procedure can be performed as follow. First, the hand image is thresholded and detected third finger and its symmetry axis such as r2 line by using wedge functions. Wedge function elements are zeros in background and ones in the object. And then it detects the centroid of the hand and identifies a vertical line

[3] https://en.wikipedia.org/wiki/Bone_age.

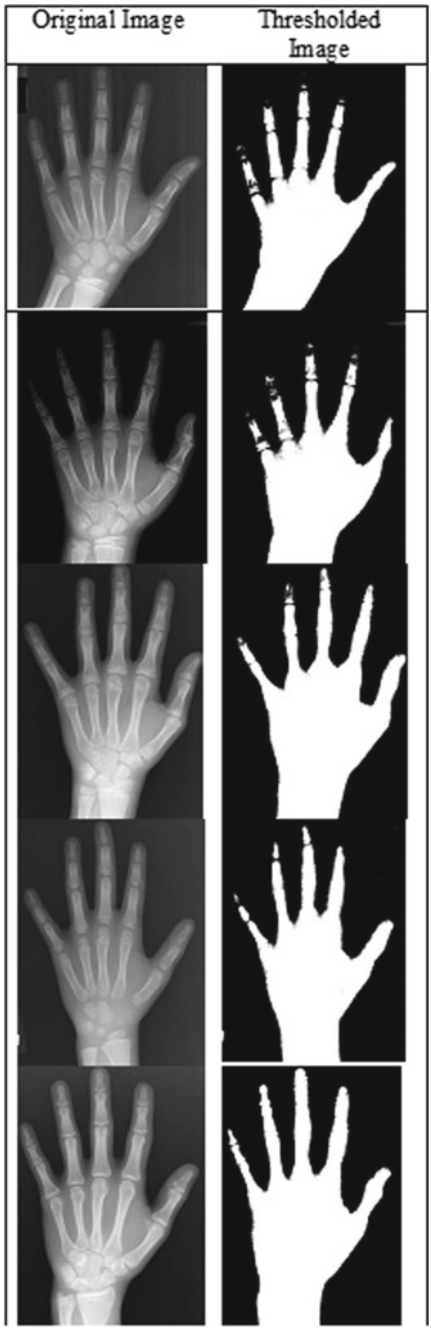


Fig. 5 Example of original images and thresholded images

passing through this centroid also called r1 line. Finally, it detects the angle between the r1 line and r2 line and rotates the image in order to do this angle is zero. If this angle is zero, the hand is oriented [20].

4 Experimental Results

In this section, CLAHE method is compared with HE and PLT methods by applying Root Mean Square Error (RMSE) and Peak Signal Noise Ratio values (PSNR) (Figs. 6 and 7).

The experimental results of CLAHE, HE and PLT methods are described in the Tables 1 and 2. The obtained results can be clearly observed in these tables. In Table 1, CLAHE method achieves 28.9857 of RMSE and 18.9579 of PSNR values in testing 20 hand images. In Table 2, CLAHE method achieves 29.074 of RMSE and 18.942895 of PSNR values in testing 40 hand images. Run time duration of CLAHE method is faster than the other methods as shown in the Table 3, Figs. 8 and 9.

According to experimental results in 20 images and 40 images, CLAHE method got better results in RMSE and PSNR values than HE and PLT methods. CLAHE's results are more effective in hand x-ray images than HE and PLT methods. Figures 6 and 7 show these results with bar chat. Figure 10 describes sample images of these methods.

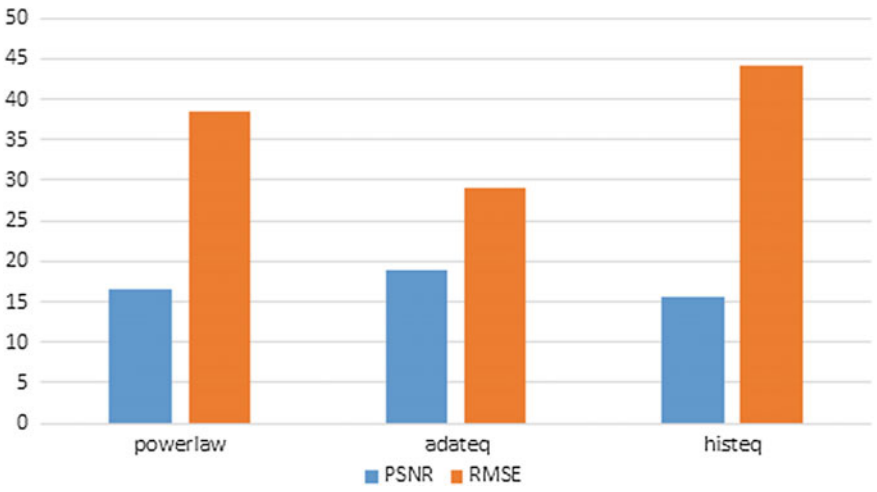


Fig. 6 The average results of 20 images

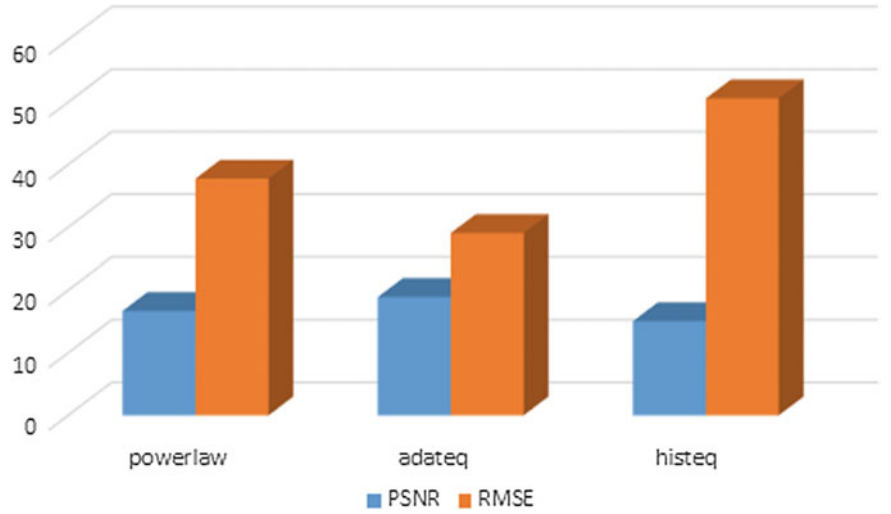


Fig. 7 The average results of 40 images

Table 1 The average results of 20 images

Method	RMSE	PSNR
Contrast Limited Adaptive Histogram Equalization (CLAHE)	28.9857	18.9579
Histogram Equalization (HE)	44.1308	15.5205
Power Law Transformation (PLT)	38.435	16.5478

Table 2 The average results of 40 images

Method	RMSE	PSNR
Contrast Limited Adaptive Histogram Equalization (CLAHE)	29.074	18.942895
Histogram Equalization (HE)	50.6187	15.0765
Power Law Transformation (PLT)	37.7848	16.71826

Table 3 Run time duration of methods

Method	Time (sec)
CLAHE	0.478132
HE	0.574086
PLT	0.572124

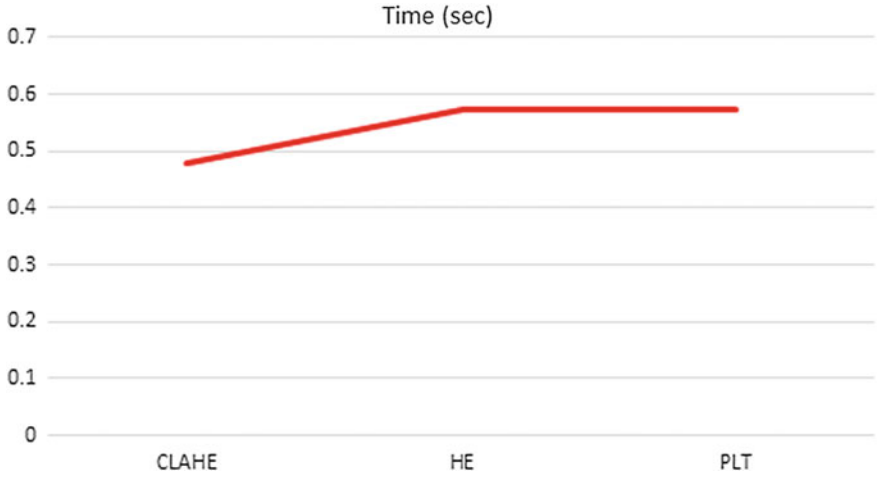


Fig. 8 Run time duration with line chart

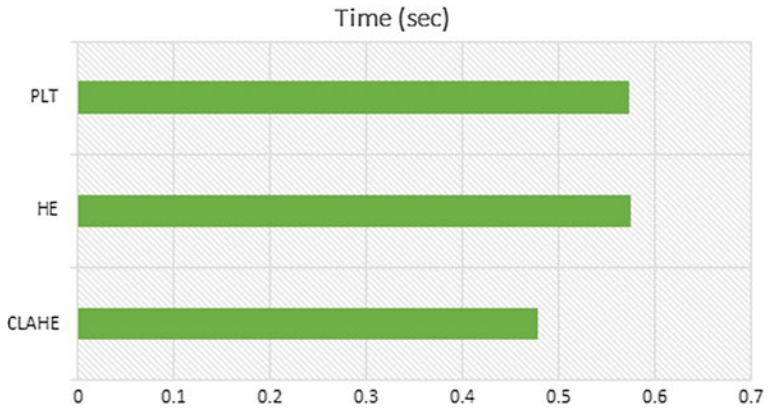


Fig. 9 Run time duration with bar chart

5 Conclusion

The system focused on the preprocessing steps of automated bone age estimation. In research work, preprocessing steps such as contrast enhancement method, background removal, and hand rotation are applied. To enhance the contrast of the hand x-ray image, the research work applied the CLAHE technique. In this paper, the contrast enhancement methods such as CLAHE method, HE method and PLT method are compared. In comparing these methods, the CLAHE method is more effective than HE and PLT methods. Running time of the CLAHE method is faster than the other methods according to the experimental results. In future work, the research work will present an automated bone age estimation by using EMROI of the middle

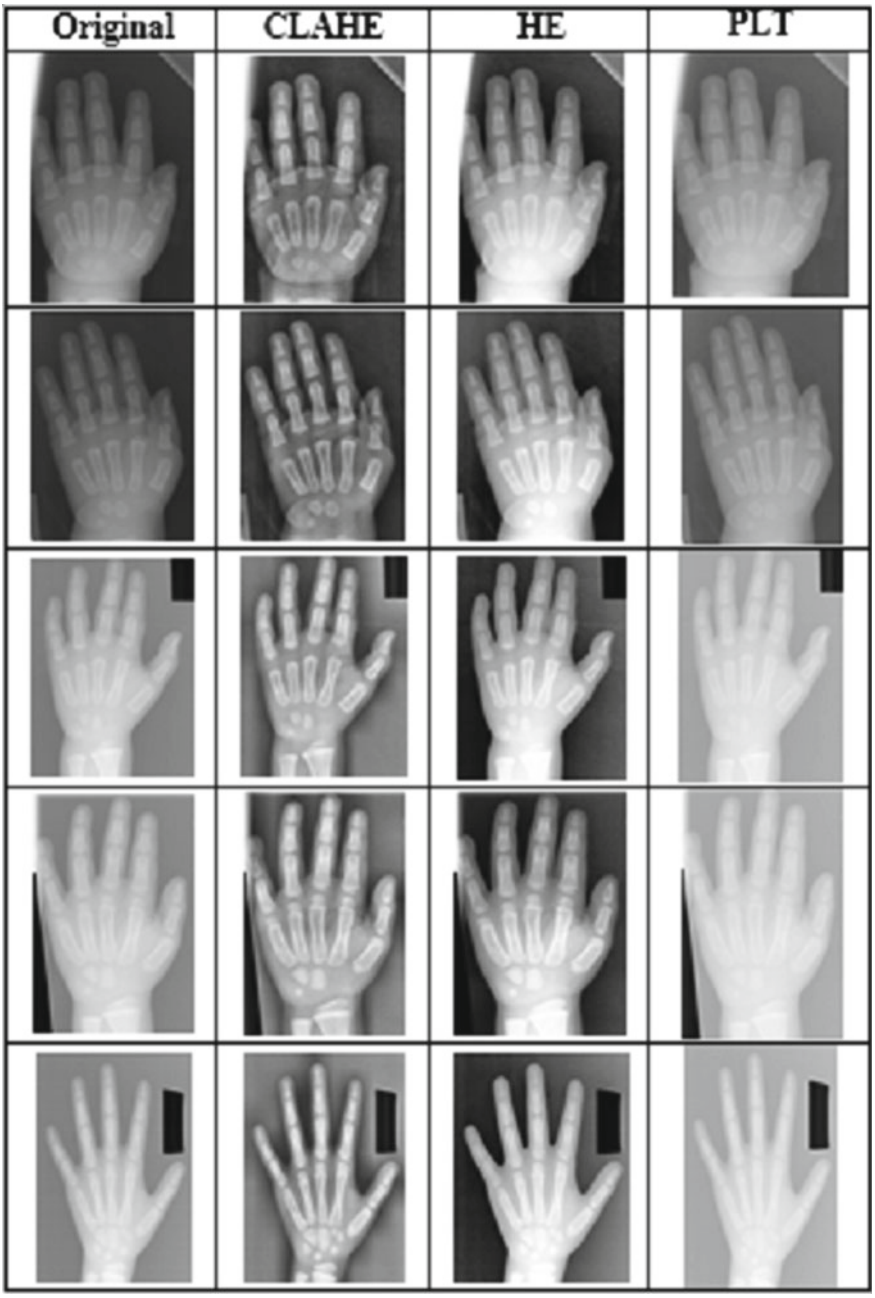


Fig. 10 The sample images of methods

finger and CROI feature analysis. The system will apply the Hidden Markov Models (HMM) classification to model different development stages of hand' bone.

References

1. Dewangan, T., Siddiqui, M.A., Bhiali, R.C.E.T.: Analysis of contrast enhancement method using modified dynamic histogram equalization
2. Thangam, P., Mahendiran, T.V., Thanushkodi, K.: Skeletal bone age assessment-research directions. J. Eng. Sci. Technol. Rev. **5**(1) (2012)
3. Lynnerup, N., Frohlich, B., Thomsen, J.L.: Assessment of age at death by microscopy: unbiased quantification of secondary osteons in femoral cross sections. Forensic Sci. Int. **159**, S100–S103 (2006)
4. Warren, M.W., Smith, K.R., Stubblefield, P.R., Martin, S.S., Walsh-Haney, H.A.: Use of radiographic atlases in a mass fatality. J. Forensic Sci. **45**(2), 467–470 (2000)
5. Spencer, R.P., Sami, S., Karimeddini, M., Sziklas, J.J., Rosenberg, R.: Role of bone scans in assessment of skeletal age. Int. J. Nuclear Med. Biol. **8**(1), 33–38 (1981)
6. Zhang, A., Gertych, A., Liu, B.J.: Automatic bone age assessment for young children from newborn to 7-year-old using carpal bones. Comput. Med. Imaging Graph. **31**(4–5), 299–310 (2007)
7. Thangam, P., Thanushkodi, K.: Skeletal bone age assessment from epiphysis/metaphysis of phalanges using Hausdorff distance. Scientific Research and Essays **7**(28), 2495–2503 (2012)
8. Güraksin, G.E., Uğuz, H., Baykan, Ö.K.: Bone age determination in young children (newborn to 6 years old) using support vector machines. Turkish J. Electr. Eng. Comput. Sci. **24**(3), 1693–1708 (2016)
9. Choi, J.A., Kim, Y.C., Min, S.J., Khil, E.K.: A simple method for bone age assessment: the capitohamate planimetry. Eur. Radiol. **28**(6), 2299–2307 (2018)
10. McNitt-Gray, M.F., Pietka, E., Huang, H.K.: Image preprocessing for a picture archiving and communication system. Investig. Radiol. **27**(7), 529–535 (1992)
11. Gertych, A., Zhang, A., Sayre, J., Pospiech-Kurkowska, S., Huang, H.K.: Bone age assessment of children using a digital hand atlas. Comput. Med. Imaging Graph. **31**(4–5), 322–331 (2007)
12. Liu, J., Qi, J., Liu, Z., Ning, Q., Luo, X.: Automatic bone age assessment based on intelligent algorithms and comparison with TW3 method. Comput. Med. Imaging Graph. **32**(8), 678–684 (2008)
13. Pietka, E., Kaabi, L., Kuo, M.L., Huang, H.K.: Feature extraction in carpal-bone analysis. IEEE Trans. Med. Imaging **12**(1), 44–49 (1993)
14. Seok, J., Hyun, B., Kasa-Vubu, J., Girard, A.: Automated classification system for bone age X-ray images. In: 2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp. 208–213. IEEE (2012)
15. Spampinato, C., Palazzo, S., Giordano, D., Aldinucci, M., Leonardi, R.: Deep learning for automated skeletal bone age assessment in X-ray images. Med. Image Anal. **36**, 41–51 (2017)
16. Giordano, D., Kavasidis, I., Spampinato, C.: Modeling skeletal bone development with hidden Markov models. Comput. Methods Programs Biomed. **124**, 138–147 (2016)
17. Giordano, D., Spampinato, C., Scarciofalo, G., Leonardi, R.: An automatic system for skeletal bone age measurement by robust processing of carpal and epiphysial/metaphysial bones. IEEE Trans. Instrum. Meas. **59**(10), 2539–2553 (2010)
18. Maitra, I.K., Nag, S., Bandyopadhyay, S.K.: Accurate breast contour detection algorithms in digital mammogram. Int. J. Comput. Appl. **25**(5), 1–13 (2011)

19. Mansourvar, M., Ismail, M.A., Herawan, T., Gopal Raj, R., Abdul Kareem, S. and Nasaruddin, F.H.: Automated bone age assessment: motivation, taxonomies, and challenges. Comput. Math. Methods Med. (2013)
20. Giordano, D., Spampinato, C., Scarciofalo, G., Leonardi, R.: Automatic skeletal bone age assessment by integrating EMROI and CROI processing. In: 2009 IEEE International Workshop on Medical Measurements and Applications, pp. 141–145. IEEE (2009)

Mis.Config: Finding Misreferred Configuration Bugs in Web Application Using Thin Slicing

Minami Yoda, Yuichi Sei, Yasuyuki Tahara and Akihiko Ohsuga

Abstract There are many web-based applications for multi-agent systems (MAS). However, developing MAS is not easy, because agents have many different variables, and it is difficult to validate them. PHP is one of the most popular web programming languages. PHP is dynamically typed, which means that the type of variable is assigned when the variable is accessed. This flexibility makes it easier to develop MAS. On the contrary, when a programmer refers variables to other PHP files, misreferences among the PHP files may occur because it is difficult to find references across files. To solve these misreferences, we propose a bug-finding tool called Mis.Config using static analysis. We used control flow graphs (CFGs) and thin slicing to realize this purpose. In our experiment, we applied our tool to real-world software to investigate whether Mis.Config can find misreferenced configurations.

Keywords Web based system development · Software engineering · PHP configuration

1 Introduction

A lot of multi-agent systems (MAS) and agent platforms have been developed for web-based applications such as [3, 4, 6, 12, 13]. A configuration file has been inevitably used for developing web-based applications and has many benefits to maintain a solid application. It gives a readability of fixed parameters that are used in applications, such as database information, the number of agents, and the behavioral

M. Yoda · Y. Sei (✉) · Y. Tahara · A. Ohsuga
University of Electro-Communications, Tokyo, Japan
e-mail: seiuny@uec.ac.jp

M. Yoda
e-mail: yoda.minami@ohsuga.is.uec.ac.jp

Y. Tahara
e-mail: tahara@uec.ac.jp

A. Ohsuga
e-mail: ohsuga@uec.ac.jp

R. Lee (ed.), *Computer and Information Science*, Studies in Computational Intelligence 849, https://doi.org/10.1007/978-3-030-25213-7_4

types of the agents. Storing this fixed information in one file gives it legibility. In particular, MAS tend to have many different parameters for controlling systems, agents, and other objects.

On the other hand, a configuration file could produce a problem. If values in the configuration file are misdefined or misreferenced, this could be a critical problem. In addition, debugging misconfiguration is time-consuming and makes developers frustrated. Developers have to check the configuration file and the source code to debug. To solve this problem, various misconfiguration finders have been proposed [2, 7, 15, 16].

PHP is a popular script programming language and dynamic typing language. PHP aims at making dynamic web sites or web systems. PHP code can be embedded into HTML files, and developers can write a system architecture and web design in the same file if they would like to.

However, this flexibility could produce an incomprehensible system. Nguyen et al. mentioned that the frequency of variable references among PHP files is high [8]. If there is a bug caused by variable references, it is hard to debug because all files related to the bug have to be checked. It is easy for developers to get confused when they have several files to debug.

Sayagh et al. studied the occurrence of multi-layer configuration options across WordPress[1] plugins and the PHP engine. This complexity often confuses developers when trying to read or understand an application's architecture [9]. Since WordPress plugins are made by individual developers, a conflict of configuration options among plugins could occur, and the applications may not work occasionally as expected. As a result, 35/484 plugins directly write one WordPress database option, and 89 plugins indirectly write one or more WordPress database configuration options. This result means that we need to access many files to detect the root cause of misconfiguration.

To aid in debugging for misreferenced configuration values across PHP files and configuration files, we propose a tool called Mis.Config, which detects misreferenced configuration values. The contributions of this research are as follows: (1) We observed that misreferenced configurations are seen in real-world PHP applications, (2) we developed a misconfiguration checker to solve misreferenced configurations by static analysis, and (3) Mis.Config is able to reproduce reported bugs from the past.

The rest of this paper is organized as follows. Section 2 presents the background of our research. Section 3 presents the design of our algorithm. Section 4 presents the results of our experiments. Section 5 discusses the related methods. Finally, Sect. 6 concludes this paper.

[1] https://wordpress.com.

2 Background

2.1 Real-World Configuration Bug

Joomla![2] is one of the most popular Content Management Systems (CMS). In Joomla 1.5, there is a bug caused by misconfiguration. This bug is located in includes/framework.php. There is `$error_reporting`, which is the configuration value that is ruled to make a flag for whether a developer wants to have an error report. In the configuration file, `$error_reporting` is initialized as "0" or "−1," and the type of these values is String.

```
var $error_reporting = '-1'; // or '0'
```

When writing the source code, the value is expected to be an Int type. In other words, when the code references the value at that line, `$error_reporting` will be compared to 0, which is an Int type.

```
$CONFIG = new JConfig();
if (@$CONFIG->error_reporting === 0) {
  error_reporting(0);
} else if (@$CONFIG->error_reporting > 0) {
  error_reporting($CONFIG->error_reporting);
  ini_set('display_errors', 1);
}
```

However, `$error_reporting` is initialized to the String type, so this value is compared to 0 (Int). The expression ===, which is the comparison operator, will be TRUE if $a is equal to $b, and they are the same type. Thus, this comparison's result will be FALSE because these variables do not match to either type or value.

Next, if `$error_reporting` does not match to 0, PHP goes to the else IF statement. In the else IF statement, this comparison will be TRUE if `$error_reporting` is greater than 0. This result will be FALSE again because either "0" or "1" do not represent a quantity.

As a result, this statement will be FALSE all the time as long as `$error_reporting` is a String type.

WebChess,[3] which is an online chess game, is published on SourceForge and has run for more than 10 years. There is a misconfiguration bug on version 1.0.0 of this software. The root cause of the bug in this file is an automated install function. WebChess 1.0.0 has install.php, which provides an automated install function.

A user can set up an application setting via this function page. This page has a form in which the user needs to input his or her preference parameters. However, when the user does not input any parameters, WebChess initializes all parameters as ' '. In the code, `$CFG_SESSIONTIMEOUT` is expected to be an Int type, but the parameter is set as ' ', which causes the error.

[2] https://www.joomla.org.

[3] https://sourceforge.net/projects/webchess/.

```
if ($_SESSION['playerID'] != -1)
{
 if (time() - $_SESSION['lastInputTime']
    >= $CFG_SESSIONTIMEOUT)
    $_SESSION['playerID'] = -1;
 else if (!isset($_GET['autoreload']))
    $_SESSION['lastInputTime'] = time();
}
```

Moreover, this reporter misunderstood the root cause of the bug. He suggested that a developer fix `$CFG_SESSIONTIMEOUT` as a String type. However, the value was used as an Int type in the code, so type changing was unnecessary. This discussion shows that the root cause for this kind of bug is difficult to diagnose, even for a skilled engineer.

RESPONSIVE filemanager[4] is a PHP application published on GitHub. We found an ambiguous configuration value in this application. An ambiguous configuration value is a configuration value that could change its type. At line 1 in the code below, `$copy_cut_max_size` is expected to be an Int type.

```
if ($copy_cut_max_size !== FALSE &&
        is_int($copy_cut_max_size)){
  if (($copy_cut_max_size * 1024 * 1024)
      < foldersize($path)){
    die(sprintf(lang_Copy_Cut_Size_Limit,
      ($_POST['sub_action'] == 'copy' ?
        lcfirst(lang_Copy) : lcfirst(lang_Cut)),
      $copy_cut_max_size));}}
```

In the configuration file, a value is set to an Int type by default. However, if developers want to ignore the limitations of the upload size, a FALSE (Bool type) is recommended to be set. Thus, the application allows `$copy_cut_max_size` to be an Int or Bool type.

Eshkevari et al. [5] revealed how PHP applications change their variable types. They said that most types of changes could be avoided by using local renaming or refactoring.

A configuration value that accepts two types would require a cost to access the configuration file and check what values are set. This also would lead to a redundancy in the source code. Simple code keeps the application easy to maintain. Thus, allowing multiple types for a variable is not recommended from a maintenance viewpoint. This option might have a possibility of bugs.

[4] http://www.responsivefilemanager.com.

Table 1 Type of configuration values

PHP apps	GitHub Star	LOC	Int	Bool	String	etc
Mural	9	950	1	2	3	0
Framadate	32	33,627	1	3	0	0
wifidog	117	65,229	0	7	26	0
Youtube-dl-WebUI	129	8,063	0	1	3	0
RESPONSIVE filemanager	397	26,114	15	26	25	0
nextcloud	407	159,828	13	30	61	0

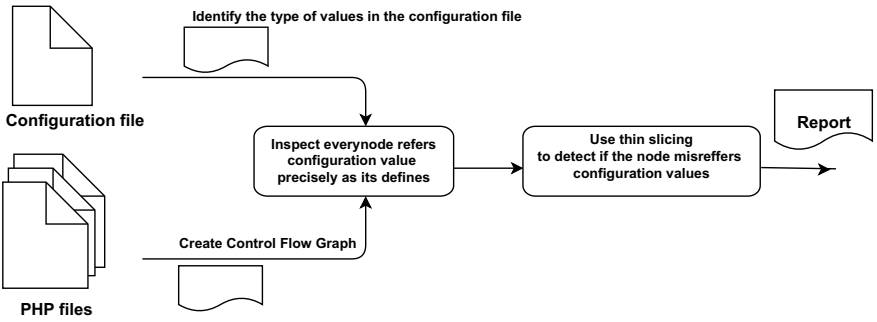


Fig. 1 Overview of Mis.Config

2.2 Commonness of Configuration Files

To make our tool Mis.Config, we observed eight PHP applications on GitHub to detect which value types are used to make configuration files. To have diversity, we chose small applications that had few stars and LOC as well as huge applications that had many stars and LOC. Table 1 shows the results. According to the results, all configuration values were expressed with any Int, Bool, or String type. As an exception, one application expressed a Bool type with "yes" or "no," but we counted this expression as a Bool type. We will focus on Int, Bool, and String types in this paper.

3 Technique

Our process was PHP static analysis using control flow graphs (CFGs) and thin slicing. CFGs are made of directed graphs. Each node has a program statement, and each edge is connected to the next node. Figure 1 shows an overview of our Mis.Config.

1. Identify the type of values in the configuration file by using regular expressions.
2. Correct the name of the configuration values and the types as pairs.
3. Create a CFG of each PHP file.
4. Inspect every node referencing configuration values precisely as they're defined.
5. Upon finding a node that references the configuration value, we inspect the node and whether this program sentence references a configuration value correctly as it's defined.
6. If the node references incorrectly a configuration value, we use thin slicing to detect variables that reference the configuration value indirectly.
7. Return the name of the misreferenced configuration value, the statement, and the name of the values that are affected by misreferenced configuration.

Algorithm 1 Algorithm of Mis.Config

```
HashMap<String, String> configList =
    getOptionsTypeList(CONFIG_PATH)
for(String filepath: List<String> phpfiles)
    CFG trace:= getControlFlowGraph(filepath)
    configList.forEach((option, type)−>
        statementMap.forEach((key,statement)−>
            if(statement.contain(option) &&
               !isWellConfigured(option,type,statement))
                result.append(option,type)
                List<String> affectedValues :=
                    getAffectedValuesList(option, trace);
                for(String value : affectedValues)
                    candidateValueList.put(option,value)
            end if
        end forEach
    end forEach
end for
```

3.1 CFGs

Allen [1] advocated for CFGs in 1970. This is a graph that explains comprehensively a route through which programs might pass. Each program is explained as a node and is connected to its relation by order of execution.

Table 2 Regular expression for extracting variable type

String	Int	Bool
".*" | '.*'	\d+	(true | false)

3.2 Thin Slicing

Slicing is a method of programming analysis proposed by Weiser [14] in 1981. This makes a criterion value and analyzes a program's flow, which can be reached from the criterion. Thanks to this method, we can obtain program statements that are affected by criterion values. The slicing proposed by Weiser is called full slicing. Sridharan [10] proposed thin slicing in 2007. Full slicing analyzes all program statements and returns the parts that are affected by criterion C. This is time-consuming because of all the analyzing. On the other hand, thin slicing only analyzes a part that is related to criterion.

Thus, this saves time. In our research, we adopted thin slicing.

3.3 Identify the Type of Values in the Configuration File

At line 1 in Algorithm 1, we extracted all configuration values of the target PHP applications to check their type and make a list of the results.

As we have discussed about the commonness of configuration variable types, we focused on an Int type, Bool type, and String type. We used regular expressions for this extraction, as shown in Table 2.

The String type is if quotemarks are used in the r-value. We defined the Int type as if the r-value is only composed of a number. The Bool type is if the r-value does not use quotemarks and is expressed by "true" or "false."

For example, if `$CFG_EXAMPLE` was defined as 123, Mis.Config decided that this value type was an Int type.

We explained an example with several configuration variables of WebChess 0.9.0. By default, `$CFG_USER`, `$CFG_NICKCHANGEALLOWED`, and `$CFG_BOARDSQUARESIZE` were defined as String, Bool, and Int types.

```
@configuration.php
    $CFG_USER = "WebChessUser";
    $CFG_NICKCHANGEALLOWED = false;
    $CFG_BOARDSQUARESIZE = 50;
```

3.4 Build CFG of Each PHP File

Next, we built CFGs of each PHP file. The getControlFlowGraph function at line 4 in Algorithm 1 makes CFGs.

We used libraries, which are produced by Nguyen et al. [8], to analyze the control flow of the PHP applications. This library works on Eclipse, which is installed PHP Development Tools. This library can analyze a reference relation among HTML, JavaScript, and PHP. Our interest was to find a reference relation in the PHP files. To realize our purpose, we ran the part of function that analyzes PHP files. These libraries provide a user interface; however, in our research, we called the library from a command line.

We implemented an original node class and CFG class to store the results. The node has information of the file name, the statement, and the variable name. We trimmed a space of the PHP statement to simplify an analysis by a regular expression. The CFG class has nodes. Nodes are connected by other nodes that follow the result of the next destination to other variables.

We gave the library a PHP file path. The library started finding a reference relation in the selected PHP file. When the library found a reference of variables, it stored information such as the position in the file, the name of the variable, the next destination to other variables, and the statement.

3.5 Inspect Every Node Referencing the Configuration Value Precisely as It's Defined

After making a configuration value list (Sect. 3.1) and creating a CFG of each PHP file (Sect. 3.2), we observed how each node in the CFG referenced configuration values at lines 4–14 in Algorithm 1. If the node referenced a configuration value in the statement (at line 7), we checked that the configuration was referenced precisely as it's defined by matching it to a configuration value list, which was created in Step 3.1 (the isWellConfigured function at line 8).

At the isWellConfigured function, we used a regular expression to detect misreferences. A rule for regular expressions is below:

Bool → **if**(!?CFG_NAME)

Bool → **if**(.+(={2}|!=){**true**|**false**}

Int → **if**(.+(>=.+|={2}\d+))

Int → .+[−+∗/]CFG_NAME[−+∗/].+

String → **mysql_connect**(w+,? w+,? w+)

This time, we focused on the IF statement. The reason we chose this regular expression is that the bugs, which are reported in real-world applications, are mostly

found in IF statements. Thus, we estimated that bugs could occur often in IF statements. If a node referenced configuration values, we checked the statement of the node with a regular expression.

To check the Int type, if the statement was a formula, Mis.Config expected the configuration value to be an Int type. In the IF statement, an Int type variable was compared to a number or other Int variables.

In the Bool type, there was a configuration value, but it was not compared to any other values in the IF statement, or there was only a configuration value and an exclamation point. In this case, Mis.Config expected the configuration value to be a Bool type. Also, if a configuration value was compared to TRUE or FALSE in the IF statement, this was also a Bool type.

In the String type, String variables in the configuration file were often used to connect databases. So we added the mysql connect function in a regular expression list. The mysql connect function is a popular PHP function for connecting to MySQL databases. The mysql connect function requires the server name, username, and password. In PHP applications, database information is often defined in the configuration file.

We introduced an example with `$CFG_NICKCHANGEALLOWED` of WebChess 0.9.0, which was mentioned in Sect. 3.3. This value was defined as a Bool type. In the code,
`$CFG_NICKCHANGEALLOWED` was referenced, as shown below:

```
@mainmenu.php
    <? if ($CFG_NICKCHANGEALLOWED) { ?>
```

In this example, Mis.Config followed the regular expression rule and detected that this value was referenced as a Bool type. Also, Mis.Config had the list of configuration variable types, which was extracted, as mentioned in Sect. 3.1. This variable type in the code and its type in the configuration list matched. Thus, Mis.Config returned that the reference of `$CFG_NICKCHANGEALLOWED` was correct.

3.6 Use Thin Slicing to Detect Local Values in PHP Files if Node Misreferences Configuration Values

In the previous phase, if the statement in the node did not reference a configuration value correctly, Mis.Config found values that were affected by the misreferenced configuration value.

We gave the name of the configuration value, which was misreferenced, and the CFG of the PHP file. A configuration value was set as a criterion of thin slicing. For example, if the misreferenced value was also stored in variable A, statements that used variable A also could be misreferenced. To warn about the effects of the misreferenced value, we ran thin slicing to find other misreferences.

Table 3 Result of Mis.Config

Case	Precision (%)	Recall (%)	F-value (%)
A	100	89	94
B	100	93	96
C	100	33	50
D	100	100	100

4 Evaluation

We conducted experiments to clarify whether Mis.Config could find misreferenced configuration values that had been inserted in applications by PHP programmers.

4.1 Condition

Four undergraduate students in information technology took the test. Their programming ability was that they could make web sites with some dynamic function, such as an e-mail form with PHP. They changed the correct configuration values into incorrect configuration values.

We experimented with Mis.Config against WebChess 0.9.0. Each student owned one WebChess application. They changed the configuration values of WebChess, or they defined new configuration values and inserted the values into a PHP file.

We did not tell students the details of Mis.Config or an algorithm in order to not influence their methods. They changed values as they wanted to and embedded those into the code.

4.2 Result

Table 3 shows the results of our experiment.

4.2.1 Case A

Student A changed the type of `$CFG_MINAUTORELOAD` to the Bool type. The value means that the min number of secs between automatic reloads of a chess game is set at a default value of 5. In the code, `$CFG_MINAUTORELOAD` was referenced as shown below.

The value was expected to be an Int type in the code because a quantity of the value was required to compare it to other values. In this case, Mis.Config decided

$CFG_MINAUTORELOAD was not suitable for this comparison because this statement required this value to be an Int type.

```
@chess.php
    if ($_SESSION['pref_autoreload']
        >= $CFG_MINAUTORELOAD)
```

$CFG_SESSIONTIMEOUT was changed from an Int type to a String type. The value means that the time out in seconds in a chess game is set at a default value of 900. In the code, the value was referenced as an Int type. Mis.Config warned that the value had been misreferenced.

```
@sessioncheck.php
    $if (time() − $_SESSION['lastInputTime'] >=
    $CFG_SESSIONTIMEOUT)
```

$CFG_MAXACTIVEGAMES was changed from a Float type to an Int type. However, Mis.Config did not warn that the value had been misreferenced. In fact, the value was not used in any codes. $CFG_MAXACTIVEGAMES was certainly defined in the configuration file but not referenced anywhere in the code.

$CFG_NEW_USER_ALLOWED was changed from a Bool type to an Int type. The value was used at index.php and was compared to TRUE. Mis.Config expected that the value was a Bool type and warned that $CFG_NEW_USER_ALLOWED should not be an Int type in this case.

```
@index.php
    if($CFG_NEW_USERS_ALLOWED==true)
```

$CFG_USEEMAILNOTIFICATION is the most used configuration value in WebChess 0.9.0. The value was changed from a Bool type to an Int type in Case A. Mis.Config expected that the value was a Bool type since this value was not compared to any value in the IF statement. Mis.Config returned that $CFG_USEEMAILNOTIFICATION misreferenced the configuration value in this case. The characteristic of a Bool type is that it does not require values to be compared as default. When a developer would like to check that the Bool value is FALSE, they add an exclamation point, '!'. Also "True" or "False" could be added per the developer's preference.

```
@chessdb.php
    if ($CFG_USEEMAILNOTIFICATION)
    if ($CFG_USEEMAILNOTIFICATION && !$isPromoting)
```

He added three new configuration values: $CFG_AAA, $CFG_BBB, and $CFG_BOARDER.

$CFG_AAA was referenced at mainmenu.php. True, the Bool type was set to this value in the configuration file. On the other hand, this value was referenced as an Int type in the code.

$CFG_AAA was used in a formula. Mis.Config warned that this value had been misreferenced.

```
@mainmenu.php
    $targetDate = date("Y−m−d", mktime(0,0,0, date('m'),
    date('d') − $CFG_AAA, date('Y')));
```

$CFG_BBB was defined as an Int type in the configuration file. In the code, $CFG_BBB was referenced as an Int type. $board[$i][$j] had a position of a chess piece, which was expected to be an Int type.

```
@gui.php
    if ($board[$i][$j] == $CFG_BBB)
```

Unfortunately $CFG_BOARDER was not used in the code.

4.2.2 Case B

In Case B, we had one false negative.
$CFG_SESSIONTIMEOUT was changed from a String type to a Float type. $CFG_SESSIONTIMEOUT was expected to be an Int type in the code. In this case, Mis.Config did not detect this misreference. The Float type is a number type, so this change is trivial compared to other types changing, such as the String type or the Bool type. However, our exception is that Mis.Config precisely detects misreferenced configuration variables. Thus, we should have distinguished the Int type from the Float type.

```
@sessioncheck.php
    $if (time() − $_SESSION['lastInputTime']
    >= $CFG_SESSIONTIMEOUT)
```

$CFG_MAXUSER was changed from an Int type to a String type. However, this value was defined in the configuration file but not used in the code.

There were two new configuration values in Case B:
$CFG MINAUTORELOA and $CFG CFG.

$CFG_MINAUTORELOA was defined as a String type and referenced at index.php. In the codes, this value was referenced as an Int type because it was compared to the number 5. Mis.Config warned that this reference was incorrect. Furthermore, as a default configuration value, $CFG_MINAUTORELOAD already existed. Mis.Config could tell a difference between $CFG_MINAUTORELOAD and $CFG_MINAUTORELOA. This shows that Mis.Config distinguishes similar value names as different values.

```
@index.php
<?php
if($CFG_MINAUTORELOA == 5){
```

```
    <title>WebChess Login</title>
}
?>
```

was defined as a String type, and the value was "cfgcfg". This variable was referenced at index.php. The echo function called $CFG_CFG. In this case, Mis.Config did not care about this reference. Because echo can call any type of variable, we think that a change of this type is trivial.

```
@index.php
<?php echo $CFG_CFG;
```

4.2.3 Case C

$CFG_NEW_USERS_ALLOWED was set as "null" in the configuration file. This value was referenced as a Bool type in the code. Null must be a critical problem if it happens at an IF statement. However, Mis.Config was not able to detect this reference. Mis.Config just aimed for the Int, String, and Bool types, so Mis.Config did not expect that the configuration values were null.

```
@index.php
    if($CFG_NEW_USERS_ALLOWED==true)
```

$CFG_DATABASE was overwritten by $CFG_SERVER. This value was used at connectdb.php. The mysql_select_dbfunction required a String type argument to set a database name. In Case C, $CFG_SERVER was a String type. Thus, this was not misreferenced. However, from the view of our tool's precision, Mis.Config did not cover the mysql_select_db function. If $CFG_DATABASE was not a String value, this misreference could not be found.

```
@connectdb.php
    mysql_select_db ($CFG_DATABASE);
```

$CFG_MAXACTIVEGAMES was changed from an Int type to a function. Also $CFG_USEMAILNOTIFICATION was changed from a Bool type to a function. Mis.Config did not analyze function type, so we could not find this misreference.

$CFG_PASSWORD was changed from an Int type to an unreadable type. This value was used at connectdb.php. In the code, $CFG_PASSWORD was expected to be a String type since the mysql connect function requires arguments to be String types. Mis.Config was not able to detect this misreference since the value (1, 1) was not expected to be a configuration value.

```
@connectdb.php
    $dbh=mysql_connect
    ($CFG_SERVER, $CFG_USER, $CFG_PASSWORD)
```

$CFG_MAILADDRESS was changed from a String type to an Int type. However, Mis.Config did not treat this as a misreference. In the code, this value was used in a character String. In a character String, any type of value can be embedded. This is why we did not treat this value.

```
@chessutils.php
    $headers.="From:WebChess<".$CFG_MAILADDRESS.">";
    $headers.="Reply−To:WebChess<"
    .$CFG_MAILADDRESS.">";
```

$CFG_BOARDSQUARESIZE was changed from an Int to a Bool type. Mis.Config did not detect this misreference. In the code, this value was used in the echo function to compose an HTML output at gui.php. $CFG_BOARDSQUARESIZE controlled the size of the image. In this case, Mis.Config did not treat a statement that accepted multiple types. However, if $CFG_BOARDSQUARESIZE did not have an invalid number, the HTML output was different from the expected one.

```
@gui.php
    echo($tmpALT.".gif' height='$CFG_BOARDSQUARESIZE'
width='$CFG_BOARDSQUARESIZE'
border='0' alt='".$tmpALT."'>");
```

$CFG_MAXUSER was changed from an Int type to a String type, as with Case B. However, Mis.Config did not warn about this misreference since this value was not used in the code.

$CFG_MAXACTIVEGAMES was also changed from an Int type to a String type. This misreference was also not detected because of it not being used in the code.

4.2.4 Case D

$CFG_NEW_USERS_ALLOWED was changed from a Bool type to an Int type. In the code, this value was referenced as a Bool type, so the misreference occurred, which Mis.Config detected.

```
@index.php
    if($CFG_NEW_USERS_ALLOWED==true)
```

$CFG_EXPIREGAME was defined as an Int type. This value represents the number of days before untouched games expire. In the code, this value was used in two parts. The first part was in a formula to determine the threshold for the oldest game permitted. The second one was in the echo function. Mis.Config detected the first one because this part is sensitive, which means that only the number type was accepted. We did not treat the second part since it was in the echo function.

```
@mainmenu.php
    $targetDate = date("Y−m−d", mktime(0,0,0, date('m'),
```

date('d') − $
CFG_EXPIREGAME, **date**('Y')));
Games will expire WITHOUT NOTICE
if a move isn't made after
<? echo ($CFG_EXPIREGAME); ?> days!

In Case D, there were two new configuration values.

In the first one, `$NEW` was defined as a Bool type. In chess.php, this was referenced as a Bool type.

@chess.php
if(isUndoing && NEW != **TRUE**)

Secondly, `$NEW2` was defined as an Int type. In chess.php, this was referenced as an Int type.

@chess.php
if(tmpIsValid && 400 == (100∗NEW2))

Finally, we tried finding the reported bug by using Mis.Config. In Sect. 2, we mentioned a bug in Joomla! 1.5. The problem of the bug was that `$error_reporting` was defined as a String type but was referenced as an Int type in the code. Because of this misreference, the IF statement in the source code did not work correctly. We treated Joomla! version 1.5.26.

Figure 2 shows a result. Mis.Config succeeded in detecting that `$error_reporting` had been misreferenced in all three PHP files.

4.3 Discussion

4.3.1 Is Mis.Config Effective in Detecting Misreferenced Bugs?

The goal of Mis.Config is finding misreferenced configuration values in PHP files. This time, we focused on a configuration reference that required a severe condition. A severe condition means that the condition only allows one type. For example, the mysql_select_db function only allows the String type.

In our experiment, we had seven new configuration values. Mis.Config detected all misreferenced new configuration values except the unused values. In spite of that, students did not know how Mis.Config works, which means that unpredictable misreferences could emerge.

This also proves that Mis.Config has an ability to find out unpredictable misreferenced configuration values. Our hypothesis that misreferenced configuration values often occur in IF statements was right.

```
1  -- includes/framework.php
2  error_reporting
3  if(@$CONFIG->error_reporting===0){
4  String, false
5
6  -- includes/framework.php
7  error_reporting
8  }elseif(@$CONFIG->error_reporting>0){
9  String, false
10
11 -- xmlrpc/includes/framework.php
12 error_reporting
13 if(@$CONFIG->error_reporting===0){
14 String, false
15
16 -- xmlrpc/includes/framework.php
17 error_reporting
18 }elseif(@$CONFIG->error_reporting>0){
19 String, false
20
21 -- administrator/includes/framework.php
22 error_reporting
23 if(@$CONFIG->error_reporting===0){
24 String, false
25
26 -- administrator/includes/framework.php
27 error_reporting
28 }elseif(@$CONFIG->error_reporting>0){
29 String, false
```

Fig. 2 Reported bug created by Mis.Config against Joomla! 1.5

4.3.2 What Kind of Configuration Value did we not Find?

In Case C, we had different types of values such as function or null. Mis.Config did not found these types of misreferences. Furthermore, modern PHP applications often use PHP data objects (PDOs)[5] to connect a database. If we were to experiment

[5] http://www.php.net/manual/en/intro.pdo.php.

with Mis.Config on a PHP application that uses PDOs, we would not be able to find misreferenced configuration values.

5 Related Work

To the best of our knowledge, our work is the first study to find misreferenced configuration values in PHP applications.

Attariyan et al. [2] proposed ConfAid, which arose from their previous work, AutoBash [11]. ConfAid is a tool that detects the root cause of misconfiguration. ConfAid statically generates CFGs from binaries not too specialized in programming language that can be treated. M. Attariyan et al. evaluated the ability of their tool to see whether it could detect the root cause of configuration of 18 real misconfigurations, such as in OpenSSH, Apache, and Postfix. ConfAid succeeded in finding the root cause in these famous applications. However, the heuristic way used to analyze is time-consuming.

Zhang et al. [16] proposed finding the root cause of misconfiguration. They focused on when a software updates its version, when the possibility of emerging bugs is high. Their tool, called ConfSuggester, checks programming statements where it is changed after being upgraded. ConfSuggester treats Java applications as bytecodes. The tool actually found the root cause of misconfiguration in largescale Java applications.

Nadi et al. [7] proposed an approach to automatically extract configuration constraints of C language. Their approach found that constraints require expert knowledge.

Xia et al. [15] proposed an automated tool using a statistical model from the natural-language description of a bug report. They experimented with their tool on 3,203 bug reports within five open-source softwares.

6 Conclusion

We proposed a tool that finds bugs caused by misreferenced configuration values in PHP applications. Many ways to find the root cause of configuration errors have been proposed. Most studies focus on applications made from Java, C, and C++ languages. These languages include static typing, so it was not necessary to care about type errors in these languages. However, PHP application bugs have been caused by type errors of configuration variables. We found bugs caused by misreferenced configuration values in a famous CMS called Joomla!. To solve these errors, we developed a tool called Mis.Config to detect whether configuration variables are referenced in the source code as it is defined. Mis.Config uses static analysis, and we used CFGs and thin slicing to realize our method.

Mis.Config was able to find bugs that were added variables and had been randomly embedded in codes. On the other hand, Mis.Config did not find the misreferenced variables of the function and null types. Moreover, because Mis.Config relies on a regular expression to check for misreferences, it is difficult to find misreferenced unprepared patterns.

As for future work, we are interested in improving Mis.Config to find errors in upgraded PHP applications.

Acknowledgements This work was supported by JSPS KAKENHI Grant Numbers JP17H04705, JP18H03229, JP18H03340, JP18K19835, JP19H04113, JP19K12107.

References

1. Allen, F.E.: Control flow analysis. **5**(7), 1–19 (1970)
2. Attariyan, M., Flinn, J.: Automating configuration troubleshooting with dynamic information flow analysis. In: Proceedings of OSDI, vol. 10, pp. 1–14 (2010)
3. Castro-Schez, J.J., Miguel, R., Vallejo, D., López-López, L.M.: A highly adaptive recommender system based on fuzzy logic for B2C e-commerce portals. Expert Syst. Appl. **38**(3), 2441–2454 (2011)
4. El Hajj, F., El Hajj, A., Chehade, R.A.: Multi-agent system vulnerability detector for a secured e-learning environment. In: Proceedings of ICDIPCl, pp. 113–118. IEEE (2016)
5. Eshkevari, L., Dos Santos, F., Cordy, J.R., Antoniol, G.: Are PHP applications ready for hack? In: Proceedings of IEEE SANER, pp. 63–72 (2015)
6. de la Hoz, E., Gimenez-Guzman, J.M., Marsa-Maestre, I., Cruz-Piris, L., Orden, D.: A distributed, multi-agent approach to reactive network resilience. In: Proceedings of AAMAS, pp. 1044–1053 (2017)
7. Nadi, S., Berger, T., Kästner, C., Czarnecki, K.: Mining configuration constraints: static analyses and empirical results. In: Proceedings of ICSE, pp. 140–151 (2014)
8. Nguyen, H.V., Kästner, C., Nguyen, T.N.: Cross-language program slicing for dynamic web applications. In: Proceedings of FSE, pp. 369–380 (2015)
9. Sayagh, M., Adams, B.: Multi-layer software configuration: empirical study on wordpress. In: Proceedings of IEEE SCAM, pp. 31–40 (2015)
10. Sridharan, M., Fink, S.J., Bodik, R.: Thin slicing. ACM SIGPLAN Notices **42**(6), 112–122 (2007)
11. Su, Y.Y., Attariyan, M., Flinn, J.: Autobash: improving configuration management with operating system causality analysis. ACM SIGOPS Oper. Syst. Rev. **41**(6), 237–250 (2007)
12. Verna, I., Bucciarelli, E., Giulioni, G., Silvestri, M.: The teaching evaluation model: a web application framework. In: Bajo, J., Escalona, M.J., Giroux, S., Hoffa-Dabrowska, P., Julián, V., Novais, P., Sánchez-Pi, N., Unland, R., Azambuja-Silveira, R. (eds.) Proceedings of PAAMS (2016)
13. Wang, M., Kawamura, T., Sei, Y., Nakagawa, H., Tahara, Y., Ohsuga, A.: Music recommender adapting implicit context using 'renso' relation among linked data. J. Inf. Process. **22**(2), 279–288 (2014). https://doi.org/10.2197/ipsjjip.22.279
14. Weiser, M.: Program slicing. In: Proceedings of ICSE, pp. 439–449 (1981)
15. Xia, X., Lo, D., Qiu, W., Wang, X., Zhou, B.: Automated configuration bug report prediction using text mining. In: Proceedings of IEEE COMPSAC, pp. 107–116 (2014)
16. Zhang, S., Ernst, M.D.: Which configuration option should i change? In: Proceedings of ICSE, pp. 152–163 (2014)

Self-Adaptation for Heterogeneous Client-Server Online Games

Satoru Yamagata, Hiroyuki Nakagawa, Yuichi Sei, Yasuyuki Tahara and Akihiko Ohsuga

Abstract Online games have been widely spread due to the increase in Internet bandwidth. In the development of an online game, we must consider the time lag problem caused by heterogeneous environments. In this paper, we apply the MAPE loop model, which consists of four key activities (monitoring, analysis, planning, and execution) for adaptation, to reduce this problem in a heterogeneous online game environment. In particular, we propose a MAPE loop control pattern for a heterogeneous client/server model. We experimentally embedded the pattern in an online game application. The results demonstrate that our MAPE loop control pattern helps to reduce the time lag problem in an online game application.

Keywords Client/server online game · Self-adaptive systems · MAPE loop control pattern · Time lag

1 Introduction

The numbers of online games and their users are rapidly growing in correspondence with an increase in Internet bandwidth. While they require real-time processing and fast interactivity, the time lag problem caused by a network communication delay

S. Yamagata · Y. Sei (✉) · Y. Tahara · A. Ohsuga
University of Electro-Communications, Tokyo, Japan
e-mail: seiuny@uec.ac.jp

S. Yamagata
e-mail: s-yamagata@ohsuga.is.uec.ac.jp

Y. Tahara
e-mail: tahara@uec.ac.jp

A. Ohsuga
e-mail: ohsuga@uec.ac.jp

H. Nakagawa
Osaka University, Osaka, Japan
e-mail: nakagawa@ist.osaka-u.ac.jp

R. Lee (ed.), *Computer and Information Science*, Studies in Computational Intelligence 849, https://doi.org/10.1007/978-3-030-25213-7_5

is one of the most serious problems. The quality of an online game decreases when a time lag occurs. In particular, match-type fighting games and first-person shooter (FPS) games require fast reaction speeds and high real-time operation accuracy. When we develop such games, mitigation of a time lag's influence on online games is an important topic to be addressed. An online game server must handle many players' commands and control their environments. Predicting these elements when constructing a heterogeneous environment, such as an online game, can be difficult. This study deals with this shortcoming by applying an adaptive method that organizes users' commands with changes in the game environments. This paper focuses on the problems, especially caused by the differences in the time lag of each client.

Online games require high permanency. Consequently, online games systems must be designed to withstand plenty of attention and changes in environment, failures and unanticipated events. A system with this kind of capability is called a self-adaptive system [1–5]. Self-adaptive systems are equipped with a certain decision mechanism and use control loops as one of their main approaches. The MAPE loop is one of these control loops and consists of four key activities (monitoring, analysis, planning and execution) to adapt to multi-agent systems [6–9]. Self-adaptive systems incorporate these components as subsystems and validate their ability to adapt. However, practical systems including these components still have many problems, in areas such as design methods, domain characteristics and the scarcity of actual adaptation examples [10].

In this paper, we propose a new MAPE loop control pattern that can ease the time lag problem by facilitating communication between a client and a server in a client-server-type online game. Our proposed pattern can prevent the substitution of an order, which occurs when a server changes an updating cycle to a gap in an order according to the difference in the clients' time lag. Moreover, we build and evaluate a simple online game application, to which we apply our proposed pattern, to show that the self-adapting ability of our proposed pattern meets the needs of client-server online games and is a possible single component of the entire process.

The paper is organized as follows. In this section, we explain the purpose of this study. In Sect. 2, we illustrate the challenges associated with a time lag in a client-server online game. In Sect. 3, we present the MAPE loop model, and in Sect. 4 we propose a MAPE loop control pattern as a method of self-adaptation in a client-server online game. We construct a self-adaptive online game based on our proposed pattern to confirm its validity in Sect. 5. In Sect. 6, we discuss the benefits observed for our proposed pattern. Finally, we present our conclusions and suggestions for future work in Sect. 7.

2 The Problem for Online Games

An online game is a video game that a number of users around the world can play through the Internet. Online games are one of the most popular entertainment applications, but they require high security, simultaneous connections from a number of users so as not to make a user feel bored and high usability. They also require

a solution to the time lag problem, to maintain the popularity of an online game. Therefore, mitigation of the influence of a time lag on online games is an important topic to be addressed. Time lag, however, is a problem that cannot be avoided by differences in heterogeneous environment such as performance of clients and geographical distances. To deal with this problem, we define the following features and characteristics as requirements of the client-server online games.

1. Environment monitoring feature
2. Configuration management feature
3. Simple configuration.

A server must dynamically determine and adapt its optimal behavior for the environmental status and for changes of clients during the runtime. Moreover, in order to guarantee that a system design can make a high runtime easy to achieve, system configuration should be simple.

2.1 The Present Condition of a Measure of the Time Lag of an Online Game

Time lags essentially occur in online games. When a time lag occurs frequently, the balance of a game collapses. In determining the cause of a time lag, distance from the network should be considered. A time lag is an important factor that can influence a user's pleasure in playing an online game. When a client joins in an online game, the client accesses a remote server. Therefore, an online game applications more or less generates a time lag. This time lag has always been an important topic in online games. Due to the time lag, the screen drawing may become extremely slow or objects in the screen suddenly move to another points. To make matters worse, large time lag causes inconsistency occurrence.

Figures 1 and 2 depict two alternative consequences of a time lag in an FPS game, a shooting video game where the player can see the game world in their perspectives. Figures 1 and 2 depict the game screens of Player A, and Player B is in the screen. When the time lag is small, if Player A attacks Player B before Player B runs away, Player A can register a hit (Fig. 1). However, in general client-server online games, the server sends processing results to all of the clients in a pre-defined time period (Fig. 3). Therefore, when there is a large difference in time lag among clients, the order of players' commands that the server receives might be different from the real order (Fig. 2). Thus, it is possible to produce unequal situations according to the time lag in which different status data to the client are generated. This fails to keep the fairness of the game.

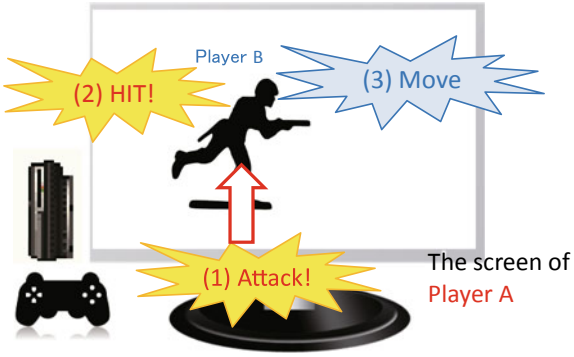


Fig. 1 Real situation. Player A successfully registered a hit

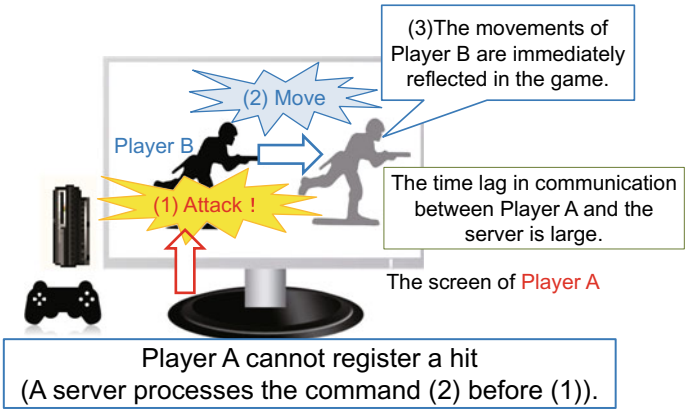


Fig. 2 Problem due to the time lag between Player A and the server. Player A's attack has not been reflected in the game yet

2.2 An Existing Countermeasure Against the Time Lag

One of the countermeasures against the time lag in an online game has already been used in some environments. Figure 4 represents a typical interaction between the server and clients in an online game. In this figure, the network delay is x (ms), the depiction of the actual processing is y (ms), and the processing time on the server side is z (ms). This interaction requires $(2x + y + z)$ (ms) to reflect the user's command to the user's screen. Figure 5, on the other hand, illustrates the countermeasure process against the time lag. This process starts to draw the screen without waiting for a response from the server. Here, if the operation is successful on the server's side, the total processing time is z (ms). This countermeasure can smoothly update the user's screen.

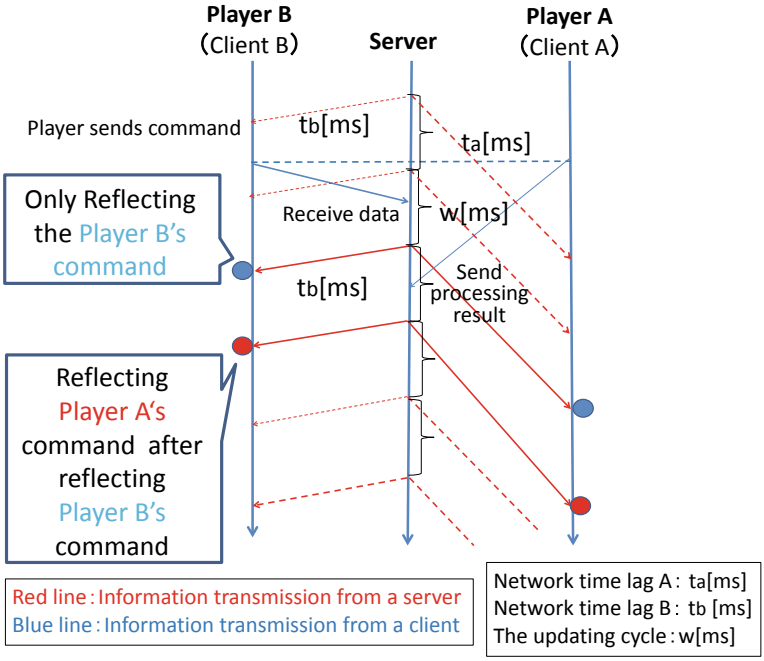


Fig. 3 Deviation caused by the time lag

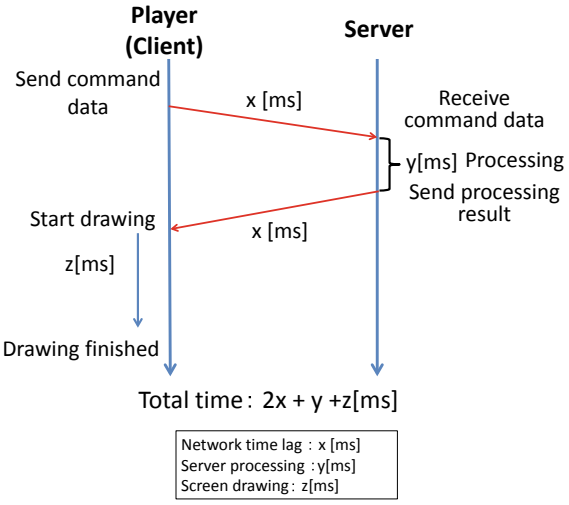


Fig. 4 No improvement for the screen display lag

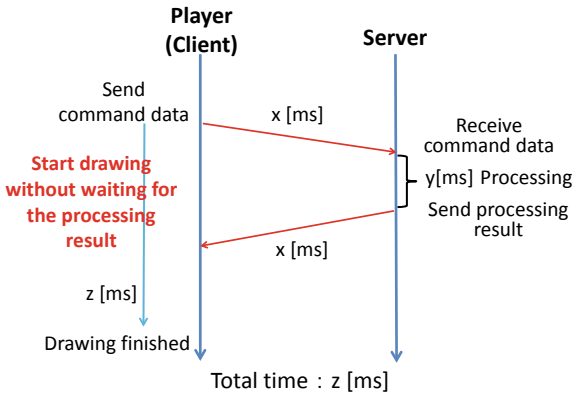


Fig. 5 Existing approach to improve the screen display lag

The countermeasure, however, suffers from the re-drawing when the server sends unexpected processing results. In these cases, it is necessary to re-draw the screen according to the received results or go back to the previous state not to cause inconsistencies. Therefore, if a game has high interactivity or multiple players, this method can not be applied.

3 MAPE Loop

The MAPE loop [11] is a mechanism of creating behaviors of a self-adaptive system [6, 7, 12]. A MAPE loop consists of four components: Monitor, Analyze, Plan, and Execute (Fig. 6).

3.1 MAPE Components

In this section, we explain each component of a MAPE loop. The Monitor component monitors the objects that affect the system performance. These objects locate within the system or the environment that the software is running [13]. Based on the information collected by the Monitor component, the Analyze component identifies

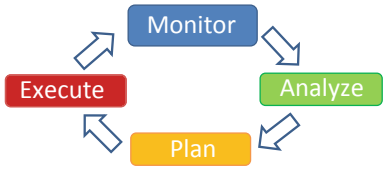


Fig. 6 MAPE loop

the current state and the elements that causes a possible failures. The component also determines whether the current state meets the requirements. Based on the results of the Analyze component, the Plan component makes plans to react to the environmental changes. The Execute component finally realizes the plans that the Plan component made.

3.2 *Existing MAPE Loop Patterns and Their Problems*

If a system is constructed over a network, the MAPE loop components must be distributed over a network to work efficiently [14]. In order to ensure the MAPE loop corresponds to the characteristics of the distributed systems, Weyns [10, 13] proposed several MAPE loop control patterns: coordinated control, information sharing, master/slave, regional planning, and hierarchical control. The coordinated control pattern is a basic structural pattern for distributed self-adapting systems. In a coordinated control pattern, the system engages in the self-adaptation process by ensuring components work with each other. In an information-sharing pattern, the system achieves self-adaptation using the Monitor components to share the acquired information. The master/slave pattern is composed of a master node and slave node. The master node contains the Analyze and Plan components, whereas the slave node has the Monitor and Execute components.

The hierarchical control pattern is hierarchically composed of MAPE loop components (Fig. 7). In Fig. 7, the M (in blue) is the Monitor component, A (in red) is the Analyze component, P (in yellow) is the plan component and E (in green) is the Execute component. Although the hierarchical pattern can be applied to vari-

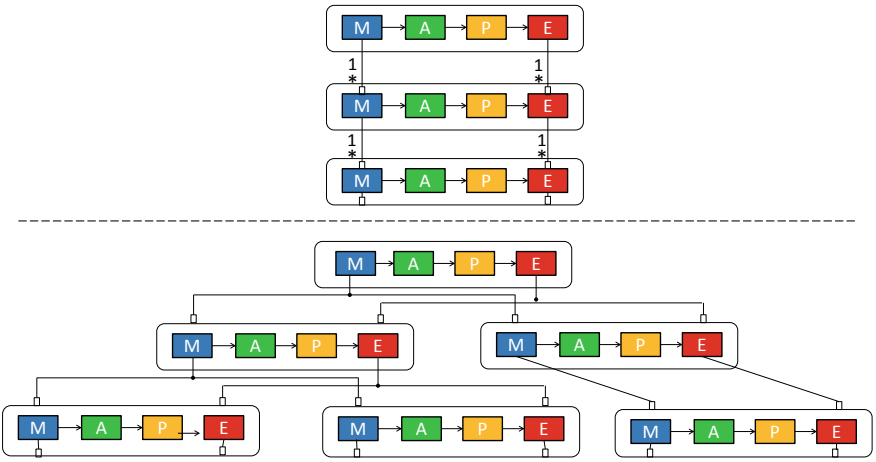


Fig. 7 Hierarchical control pattern. (Upper) general form; (Lower) an instance of the pattern

ous distributed systems, when we consider to develop an online game, it is difficult to apply it. First, since all of the nodes, i.e., the server and clients, have the full MAPE loop components, it takes extra overhead. Second, placing the full MAPE loop components to all of the clients makes these clients slowdown.

4 Proposed Pattern

To make online games adaptive, we propose a MAPE loop control pattern that facilitate the self-adaptation process of a client-server online game. We deploy the Monitor and Analyze components to the client node to monitor the client's situation and to find changes what cannot be predicted in advance. We also deploy the Plan and Execute components in the server node so that the server will be able to determine and change to its optimal behaviors immediately. Our proposed pattern has simpler MAPE loop configuration comparing with the hierarchical control pattern, so that we can expect that the real-time property will not be lost.

4.1 The Outline of a Proposed Pattern

Figure 8 shows our proposed MAPE loop control pattern. The instance diagram at the bottom of Fig. 8 shows a concrete instantiation of a pattern of four nodes of MAPE components. This pattern demonstrates a client-server application with a highly real-

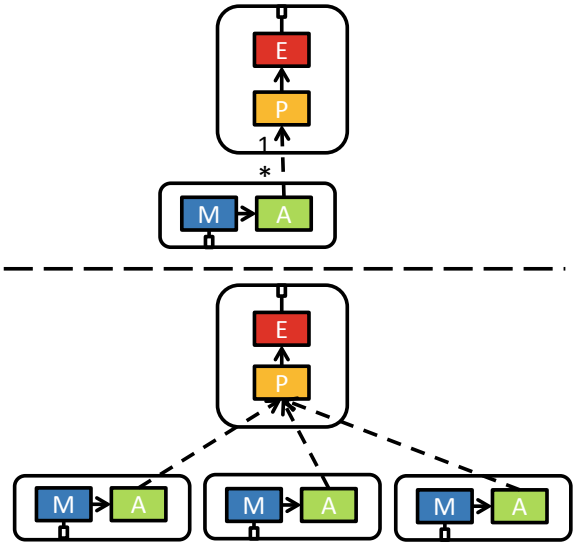


Fig. 8 Proposed MAPE loop control pattern. (Upper) general form; (Lower) an instance of the pattern that has three clients

time nature in a heterogeneous environment, aimed at a system whose server performs complicated processing. In such heterogeneous environments, it is difficult to predict changes in the environment of a client in advance. Therefore, considering time lag, deploying the Analyze component in the server node would not be desirable. This pattern is focused on the system that the server side performs complicated processing and manages multiple clients. Therefore, it is more desirable if the server grasps the analysis result of each client and changes the system's configuration and its behavior accordingly.

Considering these points, we deploy the Monitor and Analyze components at client nodes and deploy the P and E components at the server node. The Monitor component acquires the data at the client node, and the Analyze component promptly identifies the current state from the acquired data. These activities cannot be executed if they are not located at the client node. The server receives the summarized analysis results and determine what action should be executed.

4.2 *A Self-Adaptation Scenario for a Time Lag in an Online Game*

Our proposed pattern can be applied to the application that the server executes complicated process and manages all of the clients. Figure 9 shows an example that applies the pattern to the client-server online game application. When the server accurately reflects the same timing information to all of the clients, the fairness of the game is maintained. In this self-adaptation scenario, the server changes an updating cycle to conform to the communication to each client. For example, when the network delays of Client A and Client B are considerably different, the updating cycle should be

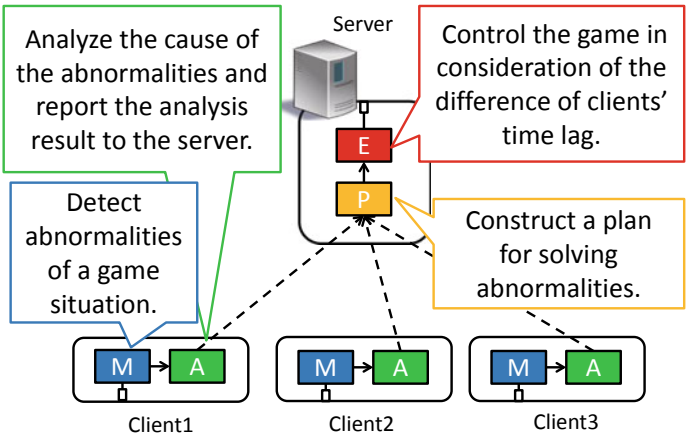


Fig. 9 An online game application using the proposed pattern

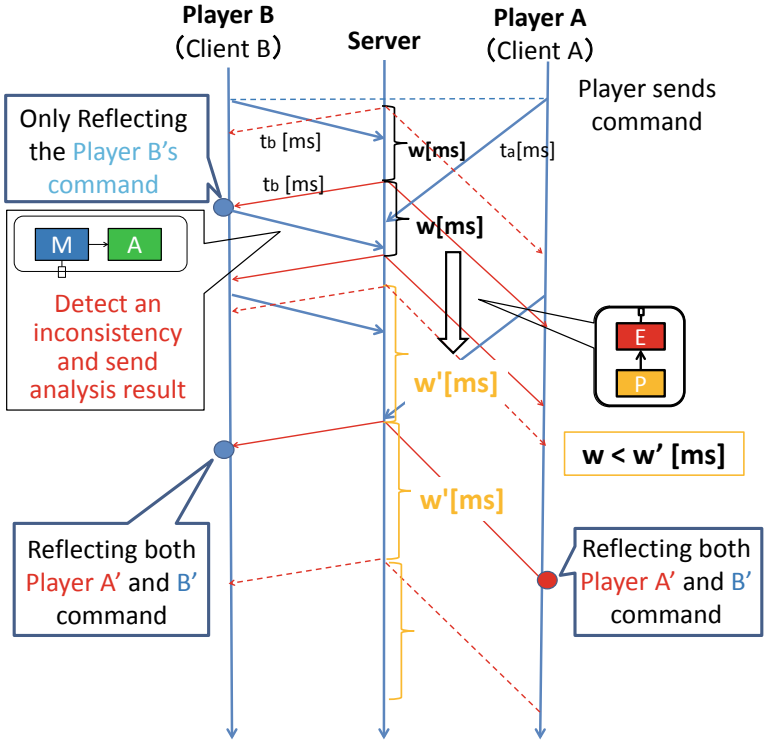


Fig. 10 A self-adaptation scenario using our MAPE loop pattern

changed from w to w' (w < w'). As a result, the server enables to obtain data from both client in the same time slot (Fig. 10).

The Monitor component detects whether an inconsistency has occurred in a client environment. Based on the information collected by the Monitor component, the Analyze component evaluates whether the time lag may cause inconsistencies. Based on the analysis result, the Plan component plans upcoming changes in an updating cycle, so that the server can receive data from all clients in the same time slot. The Execute component changes an updating cycle based on the plan generated by the Plan component. Figure 11 depicts a configuration of an online game application where the MAPE loop pattern deals with inconsistencies.

5 Evaluation

To evaluate the validity of our MAPE loop pattern, we construct a self-adaptive online game application based on our proposed pattern. The game application has a server and one or more clients (players). In this game, each player moves he/her

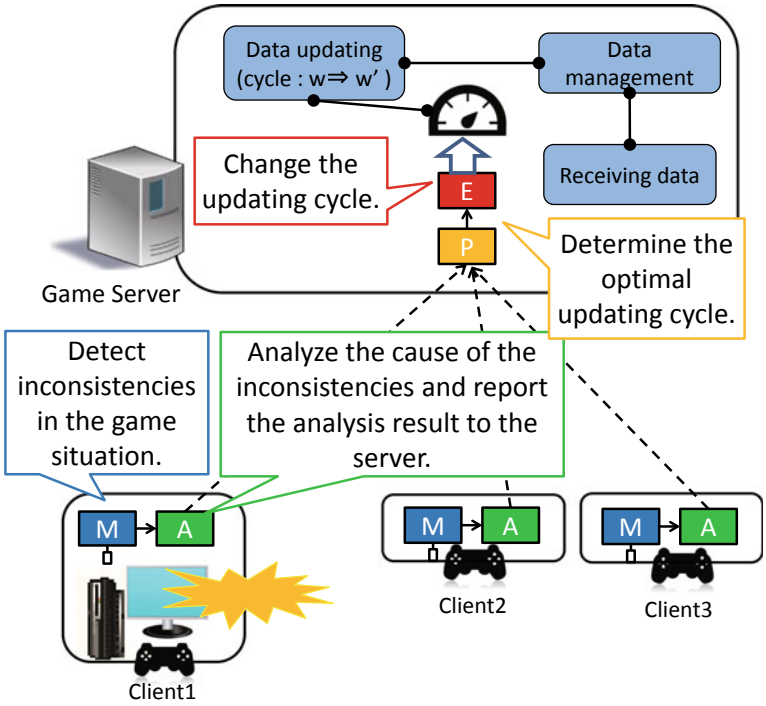


Fig. 11 A configuration of an online game application

player character in a two-dimensional map to pick up items quickly, not to be taken by other players. The server and clients interact according to the interaction diagram illustrated in Fig. 4.

We evaluated the effectiveness of our proposed pattern by counting the number of inconsistency occurrences during a game play. We defined an inconsistency as the situation where two or more clients had different data caused by the time lag.

The experiment was conducted on a virtual network using a Java application. The MAPE loop component based on this pattern was mounted in the constructed game application. In this application, each client sends the status data of the same operation to a server. The server processes the status data sent from the client and transmits the processed data to each client at a fixed period. In this experiment, the gaps between the timings according to which the client status data are reflected in the server and the differences in time lag of each client, are set as the objects of evaluation.

We conducted the experiment with two clients (Experiment 1) and with three clients (Experiment 2). We generated the time lag of each client by using a sleep time function, as shown in Table 1. To observe the behavior of the application in various environments, we prepared the following four cases:

Table 1 Differences in time lag between each client

	Client 1 (ms)	Client 2 (ms)	Client 3 (ms)
Experiment 1	0	40	N/A
Experiment 2	0	40	80

Table 2 Updating cycle

	Updating cycles (ms)	Average updating cycles (ms)
(a) Without self-adaptation	120 (fixed)	120
(b) Without self-adaptation	150 (fixed)	150
(c) With self-adaptation	from 60 to 300	120
(d) With self-adaptation	from 60 to 240	150

Table 3 Number of inconsistency occurrences in Experiment 1

Average updating cycles (ms)	Without self-adaptation	With self-adaptation
120	46 (a)	2 (b)
150	8 (c)	6 (d)

Table 4 Number of inconsistency occurrences in Experiment 2

Average updating cycles (ms)	Without self-adaptation	With self-adaptation
120	18 (a)	3 (b)
150	15 (c)	7 (d)

- *Case (a)*: The server updates the game data every 120 ms (fixed). The server does not have the self-adaptation feature.
- *Case (b)*: The server updates the game data every 150 ms (fixed). The server does not have the self-adaptation feature.
- *Case (c)*: The server updates the game data every 60 ms at first. Since the server has the self-adaptation feature based on our MAPE loop pattern, it changes the updating cycle from 60 to 300 ms. The average updating cycle is 120 ms.
- *Case (d)*: The server updates the game data every 60 ms at first. Since the server has the self-adaptation feature based on our MAPE loop pattern, it changes the updating cycle from 60 to 240 ms. The average updating cycle is 150 ms.

The updating cycle of these four cases are summarized in Table 2.

The results of Experiments 1 and 2 are listed in Tables 3 and 4, respectively. In both experiments, the server changed the updating cycle in cases (c) and (d) when one of client detected an inconsistency. Tables 3 and 4 show that the system works better with a self-adaptive feature in both experiments.

6 Discussion

This section describes our experimental results. We observed that when we added self-adaptive mechanism, which is provided by our MAPE loop pattern, to the online game, the number of inconsistency occurrences decreased. In general, larger updating cycle prevents inconsistency occurrences, that is, improves the fairness of games; however, it reduces the performance of the games. Frequent update, on the other hand, requires overloaded resources so that the game becomes suffering slowdown. This study does not take into account the optimal update period in a given online game environment. In our experiments, the servers with self-adaptive mechanisms in Case (c) and (d) simply changed the updating cycles to the maximum values when clients detected inconsistencies, and then the servers gradually accelerated the updating cycles. If we enhance the MAPE loop to learn the suitable updating cycle and timing to be changed from the history, the change operations can be more effectively executed. In terms of monitored data, the Monitor components in our experiment just collected the position data of characters and the items. In general online games, we may have to analyze the difference in time lag. Collecting various data concerning with the time, such as a communication speed between the client and the server, server load, and the specifications of client's machines, makes the self-adaptation mechanism real.

Next, we discuss the scope and potential of this pattern. In this pattern, clients monitor and analyze their environments, and the server plans and execute changes. In this pattern, the problems detected at the client side is sent to the server, and then the server deals with the problems. Some existing studies and patterns suggest the effectiveness of locating the Plan component to the central and high-performance server in cooperative distributed environments. Tajalli et al. [15] proposed a method to adapt to unexpected problems using a Roomba base station to plan a response to the problem. They locate the Plan component on the serve (base station) to enhance the performance of cleaning robots. When we deal with online games, a centralized server usually manages the game environment. From this viewpoint, locating the Execute component in the server side is reasonable. Some decentralized systems such as P2P, however, may not match with our pattern. The location of the Execute component, i.e., client side, server side, or both sides, should be decided with considering the system requirements.

7 Conclusion and Future Work

We proposed a MAPE loop control pattern for client-server online game applications. In particular, we focused on the changes of the command order due to the difference in the time lag of clients. This pattern detects the inconsistencies at the client side and deals with them at the server side. In order to show the effectiveness of this pattern, we conducted experiments on a virtual network using a Java-based online

game application. We observed that the proposed MAPE loop pattern effectively reduced inconsistency occurrences.

We plan to develop a programming framework in accordance with our pattern. Based on the framework, we will construct more sophisticated framework for online games. The current our study does not take into account the optimal updating period discovery in a given environment. If we enhance the framework to learn the suitable updating cycle and timing to be changed from the history, the time coordination can be more effectively executed.

Acknowledgements This work was supported by JSPS KAKENHI Grant Numbers JP17H04705, JP17KT0043, JP18H03229, JP18H03340, JP18K19835, JP19H04113, and JP19K12107.

References

1. Cardellini, V., Casalicchio, E., Grassi, V., Lo Presti, F.: Adaptive management of composite services under percentile-based service level agreements. In: Maglio, P.P., Weske, M., Yang, J., Fantinato, M. (eds.) Service Oriented Computing Lecture Notes in Computer Science, ICSOC 2010. LNCS, vol. 6470, pp. 381–395. Springer, Heidelberg (2010)
2. Hallsteinsen, S., Stav, E., Solbeerg, A. and Floch, J.: Using product line techniques to build adaptive systems. In: Proceedings of 10th (SPLC) (2006)
3. Hallsteinsen, S., Hinchey, M., Park, S., Schmid, K.: Dynamic software prodeuct lines. IEEE Comput. **41**(4), 93–96 (2008)
4. Morin, B., Barais, O., Jezequel, J.-M., Fleurey, F., Solberg, A.: Models@run.time to support dynamic adaption. IEEE Comput. **41**(10), 44–51 (2008)
5. Robinson, W., Purao, S.: Specifying and monitoring interaction and commitments in open business processes. IEEE Softw. **26**(2), 72–79 (2009)
6. La Iglesia, D., Gil, D., et al.: A self-adaptive multi-agent system approach for collaborative mobile learning. IEEE Trans. Learn. Technol. **8**(2), 158–172 (2015)
7. Arcaini, P., Riccobene, E., Scandurra, P.: Modeling and analyzing MAPE-K feedback loops for self-adaptation. In: Proceedings of the 10th International Symposium on Software Engineering for Adaptive and Self-Managing Systems. IEEE Press (2015)
8. Nallur, V., Cardozo, N., Clarke, S.: Clonal plasticity: a method for decentralized adaptation in multi-agent systems. In: 2016 IEEE/ACM 11th International Symposium on Software Engineering for Adaptive and Self-Managing Systems (SEAMS). IEEE, pp. 122–128 (2016)
9. He, L., et al.: Self-adaptive systems framework based on agent and search-based optimization. In: 2017 IEEE 24th International Conference on Software Analysis, Evolution and Reengineering (SANER). IEEE, pp. 557–558 (2017)
10. Weyns, D., Schmer, B., Grassi, V., Malek, S., Mirandola, R., Prehofer, C., et al.: On patterns for decentralized control in self-adaptive systems. Software Engineering for Self-Adaptive Systems II. Lecture Notes in Computer Science, vol. 7475, pp. 76–107. Springer, Berlin (2013)
11. Kephart, J.O., Chess, D.M.: The vision of autonomic computing. IEEE Comput. **36**, 41–50 (2003). IEEE Computer Society Press
12. Jacob, B., Mui, S., Pannu, J., Park, S., Raguet, H., Schneider, J., Vanel, L., Matthys, C., Bari, P., Lieurain, E., Salomon, D., Winkelbauer, L.: IBM RedBooks: on demand operating environment, managing the infrastructure (SG24-6634-01), June 2005, 2nd edn. (2005)
13. Weyns, D., Iffikhar, U., Soderlund, A.: Do external feedback loops improve the design of self-adaptive systems? A Controlled Experiment, In: Proceedings ofthe 8th International Symposium on Software Engineering for Adaptive and Self-Managing Systems (SEAMS2013) (2013)

14. Patikirikorala, T., Colma, A., Han, J., Wang, L.: A systematic survey on the design of self-adaptive software systems using control engineering approaches. In: Proceedings of the 7th International Symposium on Software Engineering for Adaptive and Self-Managing Systems (SEAMS 2012) (2012)
15. Tajalli, H., Garcia, J., Edwars, G., Medvidovic, N.: PLASMA: a plan-based layered architecture for software -driven adaptation. In: The 25th IEEE/ACM International Conference on Automated Software Engineering (ASE 2010) (2010)

A Method for Sharing Cell State for LSTM-Based Language Model

Seongik Park and Yanggon Kim

Abstract With the rapid growth in the integration of AI technology in various Natural Language Processing, the demands on the development within the fields of Natural Language Understanding and Natural Language Generation have been rapidly increasing as well. Both of these techniques analyze language, as it is naturally spoken or written by users, and must contend with a degree of ambiguity not present in formal language. For that reason, Language Modeling has been used as a key role in this area. Recently, the emerging field of deep learning, which applies complex Deep Neural Networks for machine learning tasks, has been applied to language modeling. Long-Short Term Memory (LSTM), a type of a recurrent neural network, has been adopted and has achieved reasonable results than the traditional language models. However, although LSTM-based language models have shown reasonable results by memorizing preceding cells' values, it is difficult to memorize all the information of preceding cells because they only use a 2-dimensional matrix to memorize all the information. To compensate for this limitation of memorizing problems, we propose a method for sharing cell state for a neural network-based language model, which considers all preceding cell states as a cell-stack. Our model achieved better performance compared to a traditional LSTM-based language model improving average perplexity scores from 133.88 to 124.32 for various time steps and from 141.29 to 133.62 for various hidden sizes, respectively.

Keywords Language modeling · Neural language modeling · Deep learning · Neural network · Natural language processing

S. Park (✉) · Y. Kim
Department of Computer and Information Sciences, Towson University, Towson, MD 21204, USA
e-mail: spark32@students.towson.edu

Y. Kim
e-mail: ykim@towson.edu

R. Lee (ed.), *Computer and Information Science*, Studies in Computational Intelligence 849, https://doi.org/10.1007/978-3-030-25213-7_6

1 Introduction

There has been a rapid increase in the integration of AI technology in various "smart" products, such as AI assistants in Apple's Siri and Microsoft's Cortana, and smart home products such as Amazon's Alexa and Google Home. This increase has driven demand for the development within the fields of Natural Language Understanding (NLU) and Natural Language Generation (NLG). Both of these techniques analyze language, as it is naturally spoken or written by users, and must contend with a degree of ambiguity not present in formal language. One solution to better address and understand the intricacies of natural language is through employing statistical models that learn from the features of natural language. For that reason, Language Modeling (LM) has been used as a key role in this area.

In the realm of Natural Language Processing (NLP), the process of LM seeks to build a statistical language model that approximates the probability distribution of likely subsequent words of a natural language based on the context words [1]. For example, for a sentence, the probability of the target word can be calculated as $P(W) = P(w_1, w_2, \ldots, w_n)$ based on the set of context words w, where n is the length of words prior to W in the sentence [2]. The emerging field of deep learning, which applies complex Deep Neural Networks for machine learning tasks, has been adopted to improve all aspects of statistical data analysis, including in NLP. A particular type of recurrent neural network (RNN) proven to be robust in processing sequential data is Long-Short Term Memory (LSTM) [3]. A variety of LSTM-based language models have been proposed that achieve better performance than other RNNs and traditional language models [4–7].

LSTM consists of three gates which are an input, an output and a forget gate and two states named a hidden state and a cell state. The cell state, C, is the most important part in LSTM because it memorizes preceding information and carries the information toward the cascading cells. Although LSTM has shown reasonable results by memorizing preceding cells' values, it is hard to memorize all the information of preceding cells because C is composed of only one matrix. Consequently, the values in C are squeezed as the time step becomes longer, and the history information gradually vanishes. This can cause a vanishing memory problem that can interfere with learning long-term dependency if the distance between words related to each other is long.

To compensate for the squeezing and vanishing memory problem, we propose a method for sharing the value of cell states in LSTM cells using cell state stacking method (CS-LSTM). In our model, each cell in a single network takes all preceding cell states as an input and considers those cell states as a cell-stack. However, using that stack as an input of the cell state causes a dimensional inequality problem since its dimension is different from that of input data and hidden state. Therefore, we use a convolutional neural network with a 1×1 filter to reduce the dimension of the cell stack so that the dimension of cell-stack is equal to that of the input data and hidden state. In addition, because this 1×1 filter can also be used as the forget gate, we

propose the other method to connect cell state to hidden state and input data without using forget gate.

This paper is organized as follows: Sect. 2 describes Neural Network-based language models and various types of LSTM as related work. Sections 3 and 4 present the system architecture, experiment, and discussion. Finally, Sect. 5 concludes this work.

2 Related Works

2.1 Traditional Language Model

Count-based LM is a typical type of traditional LM and is a process to construct the joint probability distribution over sequences of words. One of the popular approaches is n-gram model. According to the Markov assumption in (1), the n-gram model predicts the target word based on its n number of preceding words instead of every word preceding it in the sequence.

$$P(w_n|w_1, w_2, \ldots, w_{n-1}) \approx P(w_n|w_{n-m}, \ldots, w_{n-2}, w_{n-1}) \tag{1}$$

If the model considers only one preceding word, then it is called uni-gram LM. In the same manner, if the model considers two preceding words, then it is bi-gram LM. Tri-gram, which uses three preceding words, is most often used in practical NLP-based applications. The basic idea for n-gram LM is that we can predict the probability of w_n with its preceding n words by dividing the occurrences of the w_{n-m} to w_n by the occurrences of the w_{n-m} to w_{n-1} assuming maximum likelihood:

$$P(w_n|w_{n-m}, \ldots w_{n-2}, w_{n-1}) = \frac{count(w_{n-m}, \ldots, w_n)}{count(w_{n-m}, \ldots w_{n-2}, w_{n-1})} \tag{2}$$

However, there are disadvantages with the n-gram model. Firstly, this model relies on exact pattern. If there is a combination of words which are unseen in the training set, then the simple n-gram model will assign zero probability to that combination. This problem is called sparsity. Although back-off [8] and smoothing [9] have been proposed to compensate for this problem, this problem still remains. Another problem is that the words outside of the n-gram range will be ignored.

These limitations of transitional statistic language models have led to applying deep learning to LM for learning such syntactic and semantic features and to overcome the limitations.

2.2 Neural Network-Based Language Model

Neural Network-based language models have achieved remarkable results compared to traditional language models. One of the most important techniques which has led to the success of applying a neural network to the LM is the Recurrent Neural Network (RNN). RNN uses its internal memory (or cell) to process arbitrary input sequences. As a result, RNN can be applied to many NLP-based tasks such as speech recognition and language modeling. Currently, Long-Short Term Memory (LSTM) [3] has frequently been used instead of traditional RNN. As illustrated in Fig. 1, using LSTM allows the network to continue learning over many time steps by learning long-term dependencies using internal cell state. Many LSTM-based language models [4–7] have shown improvement over RNN-based language models.

Each LSTM cell in Fig. 2 consists of three gates which are an input, an output and a forget gate. The input gate, i, is used to determine the amount of input data that can be used, and the output gate, o, is used for deciding how much data can be used for calculating the result. Finally, the forget gate, f, decides how the amount of previous data can be forgotten. In addition to the gates, LSTM contains two states named hidden state, h, and the cell state, C. As described in Sect. 1, C is the most important part in LSTM. C memorizes preceding information and carries it toward the cascading cells. However, it is hard to memorize all the information of preceding cells because C is composed of only one matrix. Consequently, this can cause a vanishing memory problem. To compensate for the squeezing and vanishing memory problem, we propose a CS-LSTM model which forwards C_t to the subsequent cells.

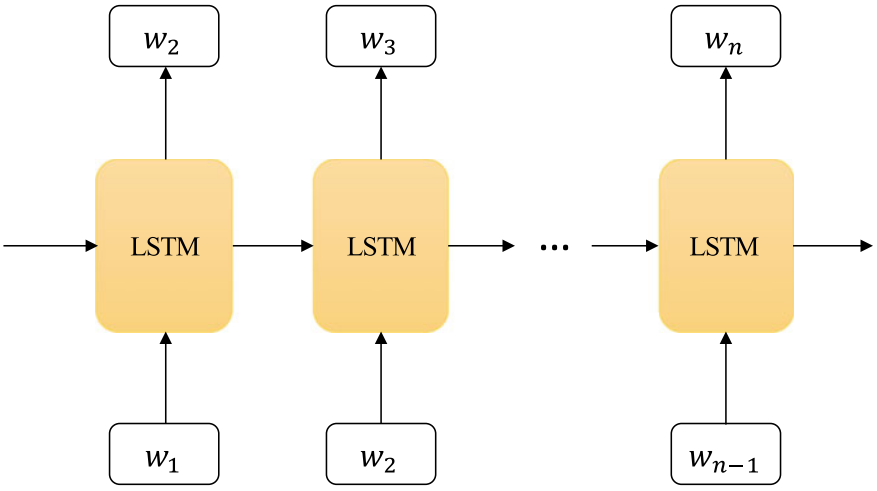


Fig. 1 Architecture of LSTM

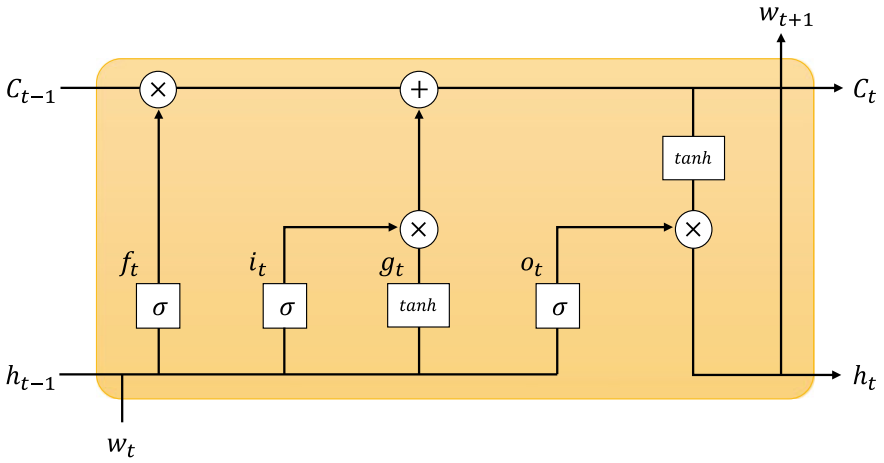


Fig. 2 Internal architecture of LSTM cell (http://colah.github.io/posts/2015-08-Understanding-LSTMs/)

2.3 Various Customized Long-Short Term Memory

There are many different types of customized LSTM. Gers et al. proposed "peephole connections," a connection to connect the cell state to the input, output, and forget gates to compensate for LSTM's limitations, which can result in information loss when the output gate value is close to zero [10]. Greff et al. suggested a coupled input-forget gate LSTM model which decides what to forget and what to add together [11]. The idea behind this is that the network can learn new data only when some older data is forgotten. By contrast, the network forgets old data only when new data is needed to be learned. Yao et al. proposed a depth gate to connect memory cells of adjacent layers [12]. The main benefit of this model is a depth gate that modulates the linear dependence of memory cells in the upper and lower layers. Even though those three approaches show the performance improvement, they still depend on a 2-D matrix to memorize all the preceding information. Thus, a squeezing and vanishing memory problem can be caused. To deal with those problems, we propose a different memorizing technique that uses a 3-D matrix to memorize every cell state value at every time step without any data loss.

3 CS-LSTM Network

As described in Sect. 1, the cell state, C, is the most important part in LSTM. C memorizes preceding information and carries the information toward the cascading cells. However, it is hard to memorize all the information of preceding cells because C is composed of only one matrix. Consequently, the values in C are squeezed as

the time step becomes longer, and the history information gradually vanishes. This can cause a vanishing memory problem that can interfere with learning long-term dependency if the distance between words related to each other is long. Therefore, in this paper, we propose a CS-LSTM model. This approach makes each cell in a single network take all preceding cell states as an input and consider those cell states as a cell-stack.

3.1 Proposed Model

Figure 3 shows an architecture of our CS-LSTM cell. A ConvNet in the figure represents a Convolutional Neural Network which performs convolutional operation with a 1×1 filter stride by 1 at a time. σ and *tanh* are non-linearity functions that denote the element-wise sigmoid $f(x) = \frac{1}{1+e^{-x}}$ and hyperbolic-tangent $f(x) = \frac{2}{1+e^{-2x}} - 1$, respectively. Finally, $\oplus$ and $\otimes$ are element-wise matrix operators.

The basic idea of our proposed model is that every cell state at time t is passed to all subsequent cells. In addition, instead of using only one two-dimensional matrix for cell state C, all Cs from the preceding cells are stacked all together in a three-dimensional matrix to store multiple Cs. This helps the network to keep memorizing all the history information without any data loss. However, using that stack as an input of the cell state causes a dimensional inequality problem since its dimension is different from that of input data and hidden state. Therefore, a convolutional neural network with a 1×1 filter is used for dimensionality reduction in order to match the dimension of cell-stack with that of input and hidden state. After that, the dimension reduced cell-stack is used to calculate a new cell state. This new architecture of CS-LSTM cell does not necessarily require the forget gate, f, because this new cell is able to memorize every C by stacking them up to compensate for the squeezing

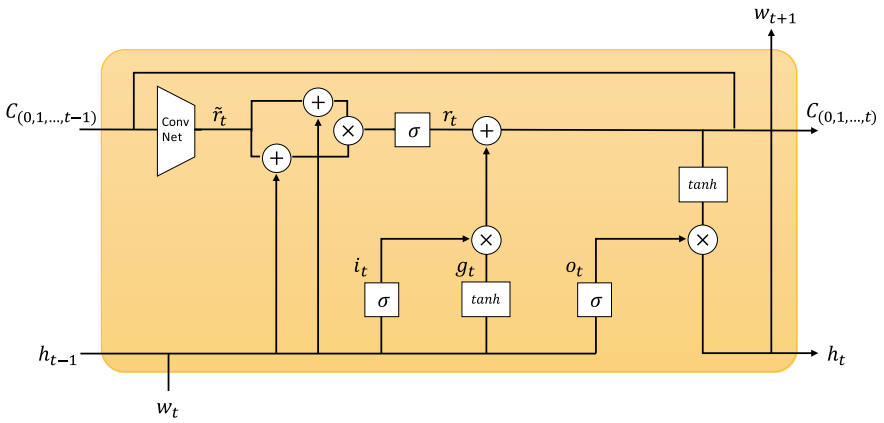


Fig. 3 Internal architecture of proposed CS-LSTM cell

and vanishing memory problem. In addition, the convolutional neural network with a 1×1 filter acts like another forget gate, learning how to reduce the values of cell-stack when reducing the dimension of the cell-stack. As a result of removing f, the connection between the input and cell state is replaced by the input w_t and the previous hidden state h_{t-1} with element-wise addition, and then two separate values are combined by element-wise multiplication to calculate a new cell state. After that, the sigmoid function is applied to composing r_t to prevent the exploding gradient problem [13]. The cell state in traditional LSTM does not contain any non-linearity function because the non-linearity can cause the vanishing gradient and exploding problem. However, the proposed LSTM is robust against the vanishing and exploding gradient problem because it stores the cell states at every time step.

3.2 Layers in CS-LSTM

3.2.1 Embedding Layer

Because the neural network only takes numerical values as its input, the word embedding, a technique to convert textual data to vector representation [2, 6, 7, 14, 15], is required. We employ the word embedding layer at the very first position of our model to transform words into vectors. Let V refer to the word vocabulary that stores every presented word in our dataset. Given a word w at time step t, w_t is converted into a vector v_n and transformed into matrix $M_n = n \times k$, where n is the size of V and k is the dimensionality of embedding.

3.2.2 Stacking Layer

The cell state in the traditional LSTM network stores the preceding information to a 2-D matrix. Compared to this, CS-LSTM stores the information to a 3-D matrix. The reason why a 3-D matrix is required is to memorize the cell states at every time step without any data loss. To keep the original cell state values, our cell has a connection to pass $\boldsymbol{C}_{(0,1,\ldots,t-1)}$ toward the end of the internal calculations and stack the original cell state values with the newly calculated cell state value. The newly calculated cell state is stacked to the 3-D matrix making $\boldsymbol{C}_{(0,1,\ldots,t)}$.

Since the cell state in LSTM has the ability to memorize the preceding information, stacking cell states at regular intervals can have a similar effect with stacking them at every time step. A skip-connection [16], a technique to skip one or more layers, is applied to CS-LSTM during the stacking process to reduce the time complexity of stacking the cell states and reducing the size of the cell-stack.

3.2.3 Convolutional Layer

Using the cell-stack as an input to the cell state causes a dimensional inequality problem since its dimension is different from that of input data and hidden state. To deal with this inequality of dimensions, a convolutional neural network with a 1×1 filter [17, 18] is applied to reduce the dimension of the cell-stack so that the dimension of cell-stack is equal to that of the input data and hidden state. With cell-stack C, a convolutional neural network performs a convolutional operation with a 1×1 filter stride by 1 at a time. The padding is unnecessary, because the convolutional neural network has the exact same output size as the input, unlike a 3×3 filter, which is the most common filter size for convolutional neural networks. The output size can be calculated by the following:

$$w = \frac{w - f + 2p}{s} + 1, \tag{3}$$

where w is the size of the width of an input, f is the size of the filter, p is the size of the padding, and s is the size of the stride. According to (3), no matter what the size of w is, it always results in the exact same size of w, since f is 1, p is 0 and s is 1. In addition, the purpose of this convolution is the dimensionality reduction, thus a pooling layer is also unnecessary. In order to compute $\widetilde{r}_t$, convolution layer sums up the contributions from C:

$$\tilde{r}_t = \sum_{x=0}^{rows} \sum_{y=0}^{columns} \sum_{z=0}^{depth} w_{x,y,z} \cdot C_{(i-x,j-y,k-z)} \tag{4}$$

This convolutional layer with a 1×1 filter reduces the dimensionality of a cell-stack into $\widetilde{r}_t$. In addition, a 1×1 filter is also trained by the network, replacing the forget gate. The output of this convolutional layer can be used as a regular cell state.

Since the forget gate is replaced with the ConvNet, new connections are required to connect input and hidden state to the cell state. To do this, element-wise addition is applied to the input and hidden state, respectively, and these results are combined into one value by element-wise multiplication. Finally, the sigmoid function limits the result from 0 to 1 to prevent a gradient explosion problem. The rest of the processes in the CS-LSTM cell is the exact same as that of the traditional LSTM cell. We will provide more details with a comparison between CS-LSTM and LSTM in the next section.

3.2.4 CS-LSTM Layer

Our proposed model does not use the forget gate f, because our model memorizes all the preceding information. Instead of using the forget gate, a new connection between input data and hidden state and new cell-stack is required. At time step t,

CS-LSTM takes an input data x_t and previous hidden state h_{t-1}, and cell-stack C_{t-1}, which contains every preceding cell state as a 3-D matrix, and calculates new hidden state h_t and cell state C_t via the following:

$$\begin{aligned} i_t &= \sigma\big(W_i \odot \big[h_{t-1}, x_t\big] + b_i\big), \\ g_t &= tanh\big(W_c \odot \big[h_{t-1}, x_t\big] + b_C\big), \\ o_t &= \sigma\big(W_o \odot \big[h_{t-1}, x_t\big] + b_o\big), \\ r_t &= \sigma\big(\big[\tilde{r}_t + h_{t-1}\big] \odot \big[\tilde{r}_t + x_t\big]\big), \\ c_t &= r_t + i_t \odot g_t, \\ h_t &= o_t \odot tanh(c_t), \end{aligned} \tag{5}$$

where W is the weight, b is the bias, h_{t-1} is the old hidden state, and r_t is the dimension reduced cell-stack. Operator $\odot$ is the element-wise multiplication, and $\sigma(*)$ and $\tanh(*)$ are non-linearity functions that denote the element-wise sigmoid $f(x) = \frac{1}{1+e^{-x}}$ and hyperbolic-tangent $f(x) = \frac{2}{1+e^{-2x}} - 1$, respectively. Once h_t is calculated, the softmax function, which converts the vector of arbitrary real values to another vector in the range 0 to 1 where the total sum is 1, calculates an output to obtain probability distribution over the next word.

4 Experiment

This section describes how to compose the dataset and how to evaluate the performance of our proposed model. To verify the efficiency of our model, we conducted this experiment by adjusting the number of hidden sizes and the number of time steps and evaluated them on the word-level perplexity and compared the results with traditional LSTM model.

4.1 DataSet

The Penn Treebank (PTB) dataset [19] is the most popular benchmark dataset to measure the quality and performance of the language models. This PTB dataset was preprocessed by Mikolov et al. [14]. They replaced rare words such as names of persons and companies with a special symbol '<unk>' and numerical words such as age and date with a symbol 'N', which are hard to be predicted by language models because of its uniqueness. As shown in Table 1, this preprocessed PTB dataset is composed of 930 k words in the training set, 74 k words in the validation set, and 82 k words in the test set, respectively. We conducted this experiment on this preprocessed PTB dataset.

Table 1 Penn Treebank corpus dataset

Data set	Size of sentence	Size of token (k)	Size of vocabulary
Training	42,068	930	10 k
Validation	3,370	74	–
Test	3,761	82	–

4.2 Evaluation and Discussion

4.2.1 Measurement and Hyperparameter

Perplexity (*PPL*) is a well-known measurement for language modeling, indicating how well a probability model predicts a sample. Given word w in a sentence S, *PPL* is calculated as:

$$\mathcal{L} = -\frac{1}{m}\sum_{i=0}^{m} \log_e P(w_i),$$
$$PPL = e^{\mathcal{L}}, \tag{6}$$

where m is the number of w in S. A lower *PPL* indicates better performance. In this experiment, we evaluated our model based on *PPL*.

Table 2 shows our experimental hyperparameters used to train each model. Based on the hyperparameters, our experiment was performed by adjusting the number of steps from 5 to 30 and the size of hidden layer from 50 to 300 to verify that our proposed model is able to achieve better performance than the traditional LSTM model.

Figure 4 shows *PPL* scores on the y-axis that change along with the various skip-connection sizes on the x-axis for test set. According to the results in Fig. 4, we set the skip-connection described in Sects. 3 as 4 which has shown the best performance. This value is based on heuristics but shows reasonable performance compared to the base CS-LSTM model which does not use the skip-connection.

Table 2 List of hyperparameters

Property	Value
Initialize scale	0.1
Number of layer	1
Batch size	20
Maximum gradient norm	5
Learning rate	1.0
Learning decay	0.5
Dropout rate	0.5

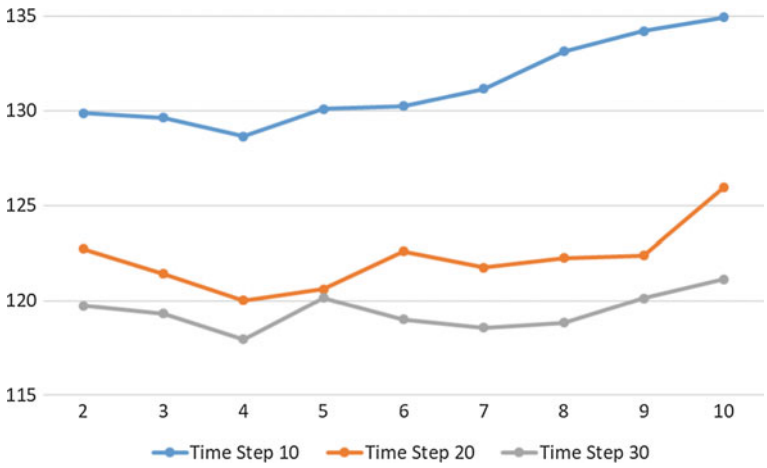


Fig. 4 Word-level perplexity for CS-LSTM model along with various time step and various skip-connection for test set

4.2.2 Evaluation and Comparison

We evaluated our approach and compared with the traditional LSTM model. Table 3 shows the performance of models based on *PPL* score. The Ours-SC in Table 3 indicates that the skip-connection applied our CS-LSTM model. As shown in Table 3, average *PPL* scores for the test set have been improved from 133.88 to 124.32 for various time steps and from 141.29 to 133.62 for various hidden sizes, respectively. Also, the skip-connection applied model shows better performance than LSTM as well and has similar performance to the CS-LSTM model. This result shows that applying our approach to the traditional LSTM based language model is able to improve the performance in terms of *PPL* with only a small amount of network size increased.

In addition to the improvement of *PPL*, we found that our model shows more robustness than the traditional LSTM with respect to the overfitting. The overfitting is one of the most difficult problems in designing and implementing the neural networks and occurs when a model biases to a training data, which can negatively impact the performance of the model on new data such as validation and test set. Figure 5 shows *PPL* scores on the y-axis that change along with the hidden size on the x-axis for test set. *PPL* value of the training set in Table 3 decreased as the hidden size increased, while *PPL* value of the test set increased explosively. This indicates that LSTM memorizes the distribution of training data too much, which is as known as the overfitting problem. In contrast, CS-LSTM shows a slight increase in *PPL* score. These results demonstrate that our model achieves relatively better performance with unseen data and is more robust for overfitting than LSTM.

Table 3 Word-level perplexity for LSTM and CS-LSTM models for various time step (above) and various hidden size (below)

Time step		Train	Validation	Test
5	LSTM	47.05	170.38	160.49
	Ours	56.43	150.82	142.14
	Ours-SC	57.02	150.99	142.87
10	LSTM	40.98	147.40	139.49
	Ours	63.48	132.69	125.65
	Ours-SC	51.17	137.66	130.16
15	LSTM	39.52	138.07	130.80
	Ours	49.12	130.96	123.93
	Ours-SC	50.44	132.17	124.83
20	LSTM	39.92	131.48	126.69
	Ours	50.05	127.36	119.80
	Ours-SC	50.27	129.87	121.47
25	LSTM	39.93	129.58	124.43
	Ours	52.58	123.48	117.69
	Ours-SC	49.67	125.28	119.54
30	LSTM	40.76	125.52	121.38
	Ours	54.44	122.04	116.69
	Ours-SC	51.00	123.97	118.20
Hidden size		Train	Validation	Test
50	LSTM	91.10	145.78	137.73
	Ours	98.58	150.13	142.81
	Ours-SC	99.48	151.81	143.83
100	LSTM	64.12	135.74	129.19
	Ours	72.34	138.58	130.39
	Ours-SC	72.63	138.93	131.83
150	LSTM	50.48	139.95	132.32
	Ours	59.02	135.68	128.49
	Ours-SC	59.79	136.17	129.46
200	LSTM	40.82	147.42	139.96
	Ours	49.90	136.83	129.65
	Ours-SC	52.42	138.77	131.24
250	LSTM	34.88	157.81	149.61
	Ours	43.84	140.37	134.04
	Ours-SC	45.83	142.09	135.10
300	LSTM	29.98	167.50	158.94
	Ours	39.00	144.59	136.32
	Ours-SC	38.75	144.42	136.16

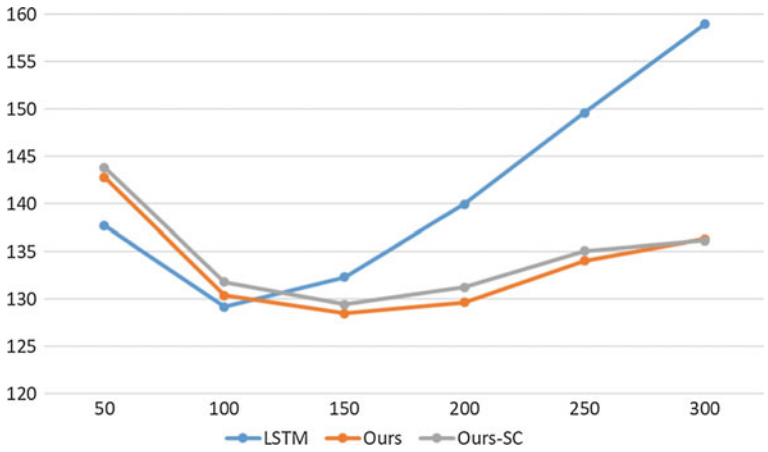


Fig. 5 Word-level perplexity for both LSTM and CS-LSTM models along with various hidden size for test set

5 Conclusion

We proposed a CS-LSTM based language model designed for dealing with the squeezing and vanishing memory problem. Our model has a new architecture of the cell that uses a 3-D matrix to store every preceding cell state making each cell in a single network take all preceding cell states as an input and considering those cell states as a cell-stack. In addition, to reduce the dimension of the cell-stack, a convolutional neural network is applied to our model. We evaluated the performance of our model on a very popular dataset and compared with the traditional LSTM network. Although our approach has not achieved the state-of-art performance, average *PPL* scores have been improved from 133.88 to 124.32 for various time steps and from 141.29 to 133.62 for various hidden sizes, respectively. In addition, our approach has shown more robustness to the overfitting problem compared to the LSTM model.

References

1. Bengio, Y., et al.: A neural probabilistic language model. J. mach. learn. res. **3**, 1137–1155 (2003)
2. Trim, C.: What is language modeling? http://trimc-nlp.blogspot.com/2013/04/language-modeling.html. Accessed 26 April 2013
3. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Comput. **9**(8), 1735–1780 (1997)
4. Sundermeyer, M., Schlüter, R., Ney, H.: LSTM neural networks for language modeling. In: Thirteenth Annual Conference of the International Speech Communication Association (2012)
5. Sutskever, I., Vinyals, O., Le, Q.V.: Sequence to sequence learning with neural networks. In: Advances in Neural Information Processing Systems, pp. 3104–3112 (2014)

6. Zaremba, W., Sutskever, I., Vinyals, O.: Recurrent neural network regularization (2014). arXiv:1409.2329
7. Zilly, J. G., Srivastava, R. K., Koutník, J., Schmidhuber, J.:. Recurrent highway networks (2016). arXiv:1607.03474
8. Kneser, R., Ney, H.: Improved backing-off for m-gram language modeling. In: 1995 International Conference on Acoustics, Speech, and Signal Processing, vol. 1, pp. 181–184. IEEE (1995)
9. Goodman, J.T.: A bit of progress in language modeling. Comput. Speech Lang. **15**(4), 403–434 (2001)
10. Gers, F.A., Schmidhuber, J.: Recurrent nets that time and count. In: Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks. IJCNN 2000. Neural Computing: New Challenges and Perspectives for the New Millennium, vol. 3, pp. 189–194. IEEE (2000)
11. Greff, K., Srivastava, R.K., Koutník, J., Steunebrink, B.R., Schmidhuber, J.: LSTM: a search space odyssey. IEEE Trans. Neural Netw. Learn. Syst. **28**(10), 2222–2232 (2017)
12. Yao, K., Cohn, T., Vylomova, K., Duh, K., Dyer, C.: Depth-gated recurrent neural networks (2015). arXiv:1508.037909
13. Bengio, Y., Simard, P., Frasconi, P.: Learning long-term dependencies with gradient descent is difficult. IEEE Trans. Neural Netw. **5**(2), 157–166 (1994)
14. Mikolov, T., Karafiát, M., Burget, L., Černocký, J., Khudanpur, S.: Recurrent neural network based language model. In: Eleventh Annual Conference of the International Speech Communication Association
15. Mikolov, T., et al.: Distributed representations of words and phrases and their compositionality. In: Advances in Neural Information Processing Systems (2013)
16. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)
17. Park, S., Song, J.H., Kim, Y.: A neural language model for multi-dimensional textual data based on CNN-LSTM network. In: 2018 19th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), pp. 212–217. IEEE (2018)
18. Lin, M., Chen, Q., Yan, S.: Network in network (2013). arXiv:1312.4400
19. Marcus, M.P., Marcinkiewicz, M.A., Santorini, B.: Building a large annotated corpus of English: The Penn Treebank. Comput. Linguist. **19**(2), 313–330 (1993)

Emotional Analysis with News Using Text Mining for Framing Theory

Jinhyuck Choi, Kwangmi Ko Kim and Yanggon Kim

Abstract Framing theory posits that the media tend to present a constructed "reality" by selecting and highlighting particular aspects of reality while obscuring or omitting others. Hence, this process leads the audience or readers to a particular understanding of reality. Depending on how each event or issue is defined, the same event can be presented and understood in different ways or through different frames. "Traditional" or typical framing research has used content analysis to identify such frames by examining major news sources that each story adopted, words used, and the tone of the stories. However, this study aims to extend this framing research to a different level by using computer-assisted analysis. This new method allows us to analyze massive data and to visualize the representation through text-mining, natural language processing, and emotion lexicon. Within the broad framing approach, this study intends to show how major newspapers of several countries depict leaders of North Korea and South Korea. The following research questions are addressed in this study: How each leader of two Koreas is represented by different countries' press? What image is dominant? How similar or different is the portrayed image? A total of eight newspapers written in English from six countries were selected for this case study: The Chosunilbo, The Korea Times, The Hankyoreh from South Korea, Uriminzokkiri from North Korea, The New York Times from USA, The Globe and Mail from Canada, People's Daily from China and Thanh Nien Daily from Vietnam. Each of these countries was selected based on the history and geopolitical relations with two Koreas. Using a keyword search such as Jong-un Kim, Guen-Hye Park, national leader, or president, we identified relevant articles from these eight newspapers and analyzed them by using text mining and Natural Language Processing (NLP). Emotion analysis is a Lexicon-based method that can detect different emotions expressed in news stories. We assumed that various degrees of emotions associated with each leader would

J. Choi (✉) · Y. Kim
Department of Computer and Information Sciences, Towson University, Towson, MD 21204, USA
e-mail: jchoi16@students.towson.edu

Y. Kim
e-mail: ykim@towson.edu

K. K. Kim
Department of Mass Communication, Towson University, Towson, MD 21204, USA
e-mail: kkim@towson.edu

R. Lee (ed.), *Computer and Information Science*, Studies in Computational Intelligence 849, https://doi.org/10.1007/978-3-030-25213-7_7

indicate the nature or orientation of reported frames. For this analysis, NRC Emotion Lexicon dictionary developed by Paul Ekman theory who identified a total of six emotions with 14,182 words was used. Then, the sentence score for each emotion was calculated. The significant contribution of this study is to present a new method of text mining and big data analysis for framing studies and to show how the overall media frames could be visualized for clear and better understanding of media representation.

Keywords Data mining · Framing theory · Natural language processing · Text mining · Emotion analysis

1 Theoretical Framework and Research Questions

The role of the media in our society has been addressed through various approaches and theories such as agenda-setting functions, framing, indexing, priming, hegemony, or concentrated media ownership structures. The research on the analysis of the media representation through a framing theory has been growing [1–13]. Framing occurs when journalists select an aspect of perceived reality and make it more important in their news story than other aspects by promoting a particular problem or interpretation. Entman [14] explains how the media frame an issue or a story [14]: Framing essentially involves selection and salience. To frame is to select some aspects of a perceived reality and make them more salient in a communicating text, in such a way as to promote a particular problem definition, causal interpretation, moral evaluation, and/or treatment recommendation for the item described. (Italics in original) (p. 52).

This framing process works by selecting and highlighting particular aspects of reality while obscuring or omitting others, a process which might lead the audience to a particular understanding of the reality. Depending on how each event or issue is defined, the same event can be presented and understood in different ways. Tankard et al. [15] described a media frame as "the central organizing idea for news content that supplies a context and suggests what the issue is through the use of selection, emphasis, exclusion, and elaboration" (p. 3) [15].

In an effort to establish a cohesive model of a framing theory, Scheufele [16] proposed a process model of framing by identifying four key processes: frame building, frame setting, individual-less processes of framing, and a feedback loop from audiences to journalists [16]. Three different research directions can be noted under a "frame building" process: The first, a simple and "basic" approach, is to show how news organizations frame a certain event or issue to accord with their existing policy or official line. The second is to show how the media portrayed the similar events differently to support its own government's interest or international relations. The third is to indicate how different countries report the same event differently by adopting different frames to fit their perspectives and circumstances.

Within the first direction, Kim [12] examined how the U.S. press covered its cigarette exports policy to Asia, and presented that this policy was framed and pre-

sented as a trade issue to support its official government position [17]. This study was limited in a sense that it did not offer any comparative analysis.

The studies by Entman [4] and by Kim [18] illustrate the second research direction. Entman [4] highlighted how two leading newsmagazines, Time and Newsweek, framed similar flight downing incidents differently through their headlines, the content and size of the cover page photos, and graphics [4]. According to his study, these two U.S. newsmagazines portrayed the Soviet military's downing of a KAL flight in 1983 as a moral outrage, while presenting the American military's downing of an Iran Air flight in 1988 as a technical problem. Kim [18] also suggested that the U.S. media defined two similar student democratic movements (in Korea in 1980 and in China in 1989) in a manner that coincided with the U.S. government's interests and foreign policy [18]. The Chinese student movement in 1989 was positively viewed as a democratic movement to protest against the communist Chinese government, whereas the Korean student movement was negatively presented as a rebellious, peace-threatening movement against the "well-established" Korean government.

The third research group is formed with recent studies that compared news coverage of a certain issue/event by different countries' news organizations. In a comparative analysis of the news coverage of the Iraq war by five Asian countries (India, Sri Lanka, the Philippines, Pakistan, and Indonesia), Maslog et al. [19] showed that the news coverage of this issue was framed differently by each country's perspective [19]. Overall, news stories were framed as peace journalism or "neutral." However, each country's religion and news sources were two important factors that shaped the coverage differently. For example, non-Muslim countries (e.g., India and Sri Lanka) adopted more war journalism framing than the Muslim countries (e.g., Pakistan and Indonesia), and viewed the Iraq war as a justified means of eliminating terrorism. Tian and Stewart [20] also revealed that two different countries' media can not only report certain aspects of the same event similarly but also frame and present it differently using several factors [20]. Their comparative analysis of the SARS crisis between CNN and BBC online texts showed that both news stories focused on the SARS effects on public health and framed the issues from a global perspective. However, it also indicated that CNN and BBC framed the SARS crisis in different ways by its own cultural, geographical, and economic factors. For example, the economic implications and the control issue of this outbreak were more prominent in the CNN news than in the BBC news. Also, due to close military and economic relationships with Taiwan, the CNN reported more stories on the role of Taiwan in the SARS crisis than the BBC. Similarly, Toronto received a higher level of attention in the CNN reports than in the BBC reports due to the fact that the U.S. and Canada share national borders and are large trading partners. On the other hand, the U.K. does not have such relations with Toronto, which is why the BBC treated Toronto as one of the SARS-affected areas, such as China.

The power of framing is especially strong in news reports of foreign events or foreign policy in which most audiences do not have direct experience or knowledge. "This is why," Entman argues, "exclusion of interpretations by frames is as significant to outcomes as inclusion" (p. 54) [14]. "Traditional" or typical framing research has conducted content analysis to examine major news sources that each story adopted,

and identify frames used in the news coverage. As major news sources, journalists often depend on government officials, congressmen, business executives, and public relations professionals as authorities, experts, or "objective" voices for information [9, 21].

However, this study aims to extend this framing research to a different level by taking interdisciplinary approach through computer-assisted analysis. Within the broad framing approach, this study intends to show how major newspapers of several countries depict leaders of North Korea and South Korea. The following research questions are addressed in this study: How each leader of two Koreas is represented by different countries' press? What image is dominant? How similar or different is the portrayed image of each leader of two Koreas appeared by various countries' media?

2 Related Work

2.1 Natural Language Processing

In order to analyze text data, we used a natural language processing of "Stanford Core NLP" made by The Stanford Natural Language Processing Group [22]. This methodology gives us the opportunity to refine text data we want such as Part of Speech (POS), Lemmatization, Sentence classification, and Tokenization.

In the English languages, words are categorized into the 8 major parts of speech.

In order to refine the data, it is important to grasp part of speech because it is necessary to delete unnecessary words under certain circumstances. For example, the word 'Like' can be completely different depending on the part of speech.

Although we can see several words of similar form in English, words have the origin of words. For example, 'is', 'are', 'were', and 'was' could be converted to 'be'. By doing this, we can reduce the time and increase the accuracy of data analysis by refining a large amount of data.

Stop words are natural language words including very little meaning for analyzing such as 'and', 'the', 'a', 'an', and similar words or this also includes cases where it cannot be cleared by refining operation such as special symbols.

2.2 Emotion Analysis

The main goal of emotion analysis is to find people's thoughts and interests through an unspecified number of data sets. This idea requires emotion analysis that can be analyzed in more detail than traditional sentiment analysis. There are several sets of emotion data, but this research used NRC Emotion Lexicon [23].

2.3 K-Means

K-Means is one of the simplest unsupervised machine learning methods and is an algorithm that is optimized for classifying each data values [24]. This is a method used in clustering analysis in data mining and statistics. We can use this algorithm to perform grouping operations. For example, we can group several people into similar tendencies.

3 Method

By March 25, 2017, 49.6% of the world's population is using the Internet [25]. Thus, the world population watching the news using the internet is increasing, and news from around the world are provided in English. So, we built a system to gather a thought that we considered influential (Fig. 1).

Instead of using traditional content analysis, the main goal of this study is to explore the ideas of each country about the same issue using data mining and big data technology. We collected 8 news from 6 countries. The six countries are S. Korea, N. Korea, United States, China, Canada, and Vietnam. We choose three influential news in South Korea. These are progressive, conservative and moderately oriented. The rest of the news consisted of news that is considered influential in each country. The list of news is Chosun, Korea Times, Hankyoreh, Globe and Mail, People China, Uriminzokkiri, Thanhnien, and Newyork Times [26–33].

In order to apply framing theory, it is chosen as articles of North Korea and South Korea news. North Korea and South Korea as the first filtering keywords and

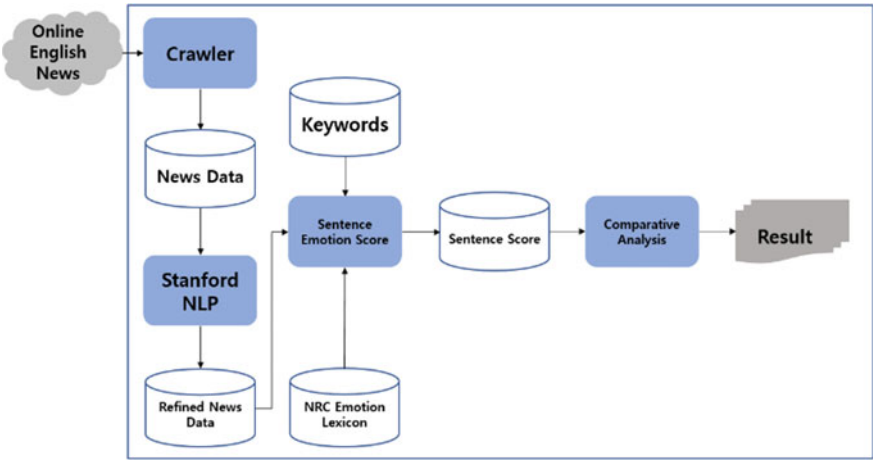


Fig. 1 Comparative analysis model of news

the second filtering were based on the names of the president and supreme ruler in Table 1. The articles can be extracted from the criteria for the second filtering keywords based on 50 articles from the latest news in each country.

The whole system is divided into four parts which are Crawler, NLP (Natural Language Processing), Sentence Emotion Score, and Comparative Analysis. The Crawler has the ability to collect real-time and previous data based on selected keywords. Second, NLP (Natural Language Processing) is a natural language processing method based on collected data. Third, the Sentence Emotion Score section is to score sentences based on the NRC Emotion Lexicon [23]. Lastly, Comparative Analysis is for a visualization based on the Sentence Emotion Score. In conclusion, we analyze the sensitivity of each news through this system. We compare the characteristics of each news about the same social issue. We will further explain each system part in the detail below.

The first part of the whole system is the Crawler. The main goal of Crawler is to gather real-time or previous articles. JSOUP has an HTML parser function [34]. This means that we can gather news in real time based on keywords and also collect previous news on Table 1. So, we systemized the Crawler part on the whole system based on JSOUP. However, because New York Times News was encrypted, it uses the API to collect articles related to keywords in the latest order [35].

The second is the NLP (Natural Language Processing). This system works to refine news data collected through Crawler [36]. As the first step, we used Stanford NLP, Sentence Parser, to cut each article by sentence because we wanted to identify the words in the sentence including the keywords [37]. After that, sentences are removed from unnecessary words called by Stop word. For example, words that do not have emotional elements such as non-verbs and prepositions are removed [38]. Lemmatization was performed thereafter, because of the grammatical problems of English such as 'organize', 'organizes', and 'organizing' which are families of derivationally related words with similar meanings [39]. In conclusion, through the NLP system, we refined based on the collected news data, the sentence to be confirmed based on the keywords. Unnecessary words were removed based on the above process.

The third part is the Sentence Emotion Score system. The main goal in this section is scored based on news data refined through NLP. Traditionally, data mining and text mining are primarily sentiment analysis. However, the results of most sentiment analysis are positive or negative. We have examined the emotions felt by humans for more accurate emotional analysis, Paul Ekman said that human emotions can be classified into six which are Joy, Sadness, Fear, Anger, Surprise, and Disgust [40]. NRC Emotion Lexicon is a dictionary constructed in the way suggested by

Table 1 List of filtering keywords

Categories	Keywords
North Korea	Kim Jong-un, Kim Jong Un, Jong-Un, Kim Junior, Jong Un, Leader & Jong-Un, Leader & Jong Un
South Korea	Park Geun-hye, Geun-hye, Geun-hye, Geun hye, President Park

Paul Ekman, so we used this dictionary to calculate the Sentence Score [41]. The dictionary consists of 14,182 words.

In this dictionary, we used a total of six emotions, and as shown in Table 2, if there is an emotion in the word, the number is expressed as 1. If there is no emotion, the number is expressed as zero. If we have words that are sensitive to the particular word we are looking for, we present a new method of summing each emotion. For example, we can select a specific word, Kim Jong-un, and extract sentences related to that word. Thus, after identifying words of emotion for Kim Jong-un, we can check sentence scores for emotions. As Table 4 shows, if we assign keywords to Kim Jong-un, we extract only the sentences related to Kim Jong-un, and extract words with good and bad feelings in it. As shown in Table 2, Bad has 1 point for Anger, Disgust, Fear, and Sadness and Good has 1 point for Joy and Surprise. This is how all sentences are scored (Table 3).

Finally, we systemized Comparative Analysis. We get a numerical value comparing each news based on Sentence Emotion Score. The number of each news is different and the document length is also different though out the same period. Therefore we could see the different emotions of each news by dividing it based on the total motion of words.

In a nutshell, if we want to get the sentence score for Kim Jong-un, the system finds the sentence for Kim Jong-un as Table 4, and then the system searches for the words contained in the NRC Lexicon as shown in Table 2, and then the system continues to add scores for each emotion. As mentioned above, the sum of the six emotion values is the denominator. The value of each emotion is located in the molecule. Thus, we compute the percentage of each emotion by comparing it with the total emotion values.

To conclude, we constructed a system for collecting articles to analyze the news. Based on this, we proposed a method and system for analyzing each news through natural language processing and scoring sentence, and each numerical data is repre-

Table 2 Example of NRC Emotion Lexicon 1

Word	Anger	Disgust	Fear	Joy	Sadness	Surprise
Bad	1	1	1	0	1	0
Good	0	0	0	1	0	1

Table 3 Example of NRC Emotion Lexicon 2

Emotion	Word
Anger	abuse, aggressive, angry, annoy, arson, assail, …
Disgust	aberration, abhor, abnormal, abuse, affront, …
Fear	abandon, abhor, absence, abuse, abyss, accident, …
Joy	absolution, accolade, achieve, acrobat, adore,
Sadness	abandon, abduction, abyss, affront, ail, …
Surprise	abrupt, alarm, alerts, allure, ambush, …

Table 4 Method of sentence calculation

Keyword	Sentence
Kim Jong-Un	Kim Jong-un is good and bad. Park Geun-hye is good and bad

sented as radial graphs for visualization, so we can identify the differences in each news using graphs, then we can cluster each news using K-Means algorithm using the results of each news.

4 Experiment

We compared the news about Kim Jong-un, the leader of North Korea, and Park Geun-Hye, the president of South Korea. We will divide each part and explain the result.

4.1 North Korea(Kim Jong-Un)

The news collection period for Kim Jong-un is from April 11, 2012 to February 25, 2017. On April 11, 2012, Kim Jong-Un became Secretary of the 1st Labor Party (Fig. 2).

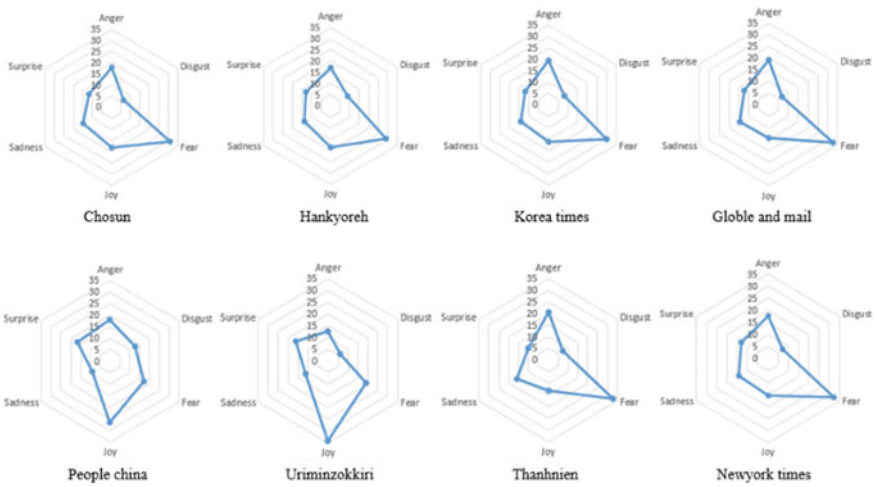


Fig. 2 Radial charts of Kim

Table 5 Result of North Korea and Kim articles

News	Number of articles	Number of Kim's sentences
Chosun	3,910	1,430
Hankyoreh	5,450	1,137
Korea Times	2,902	1,607
Globle and Mail	1,391	351
People China	793	139
Uriminzokkiri	960	697
Thanhnien	547	278
Newyork Times	2,026	626
SUM	17,979	6,265

As Table 5 shows, we collected 17,979 articles during the period mentioned above, and the total number of sentences is 462,968. There are a total of 6,265 sentences including Kim Jong-un.

We extracted the result of each emotion based on the word with emotion as shown in Table 6. The Uriminzokkiri has the least words related to Anger and the Thanhnien has the most words. A Chinese newspaper, People China, has the highest percentage about Disgust and Uriminzokkiri has the lowest percentage. In terms of Fear, Vietnam newspaper, Thanhnien, has the highest percentage and Chinese newspaper, People China, is the lowest percentage. In terms of Joy, the North Korean newspaper showed the highest percentage and Vietnam newspaper showed the lowest percentage. In terms of words including Sadness, the Vietnamese media has the highest figures and the Chinese media has the lowest figures. In terms of Surprise, the Chinese media has the highest number of results. Based on this result, we can see it as a radial chart difference.

Table 6 Emotion result of N. Korea and Kim

News	Anger (%)	Disgust (%)	Fear (%)	Joy (%)	Sadness (%)	Surprise (%)
1	17.94	6.66	30.81	18.01	14.69	11.85
2	17.03	8.89	29.15	18.4	13.81	12.7
3	19.47	8.06	29.96	16.13	14.41	11.94
4	18.93	6.92	32.72	14.27	14.77	12.36
5	18.05	12.5	17.36	26.38	9.02	16.66
6	12.64	6.15	19.18	34.39	11.31	16.31
7	20.7	7.50	32.95	12.91	15.85	10.06
8	17.55	7.28	32.11	15.53	14.36	13.13

Chosun = 1, Hankyoregh = 2, Korea times = 3, Globle and mail = 4, People china = 5, Uriminzokkiri = 6, Thanhnien = 7, Newyork times = 8

In conclusion, if we look at the radar charts, we could see that Chinese and North Korean newspapers are different from other newspapers. For objective analysis, we clustered each news through K-means technique mentioned above (Fig. 3).

In Table 7, the K-Value meaning is a value that determines how many news are to be clustered in the same group [24]. When the value of K is 2, it can be seen that North Korea news, Uriminzokkiri, and Chinese news, People China, are similar because the two newspapers are clustered into 2. The rest of the news is shown by 1. When the value of K is 3, it can be seen that S. Korea, Canada, and the United State are similar because the newspapers are clustered into 3 and Vietnam is unique as 2.

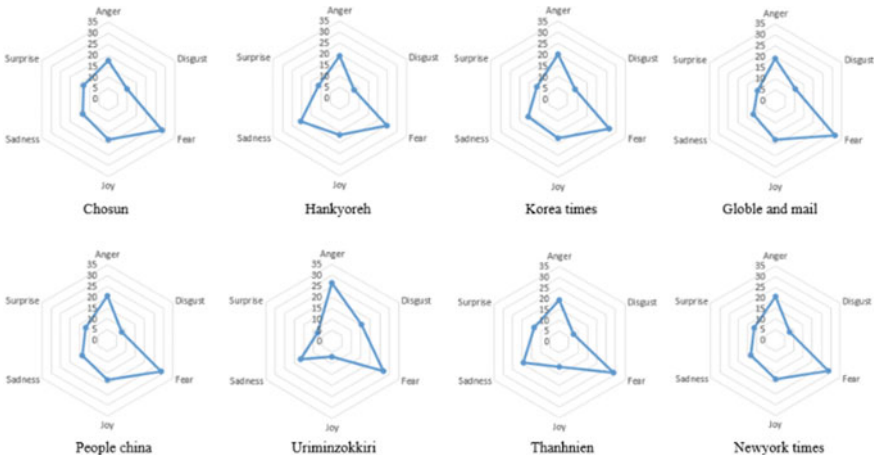


Fig. 3 Radial charts of Park

Table 7 K-means results of N. Korea and Kim

News	K-value 2	K-value 3
Chosun	1	3
Korea Times	1	3
Globe and Mail	1	3
People China	2	1
Uriminzokkiri	2	1
Hankyoreh	1	3
Thanhnien	1	2
Newyork Times	1	3

Table 8 Result of South Korea and Park articles

News	Number of articles	Number of Kim's sentences
Chosun	1,173	1,246
Hankyoreh	2,267	4,320
Korea Times	9,509	15,485
Globle and Mail	1,894	677
People China	3,186	568
Uriminzokkiri	343	386
Thanhnien	840	140
Newyork Times	1,931	232
SUM	21,143	23,054

4.2 South Korea (Park Geun-Hye)

The news collection period for Park Geun-hye is from February 25, 2013 to March 10, 2017. February 25, 2013 is the time when the Park Geun-hye government was established. On March 10, 2017, Park Geun-hye was impeached and resigned from the presidency. As Table 8 shows, we collected 21,143 articles during the period mentioned above, and the total number of sentences is 529,220. There are a total of 23,054 sentences including Park Geun-hye.

We extracted the result of each emotion based on the word with emotion as shown in Table 9. The Chosun has the least words related to Anger and the Uriminzokkiri has the most words. North Korea newspaper, Uriminzokkiri, has the highest percentage of Disgust and Thanhnien has the lowest percentage. In terms of Fear, Canada newspaper, Globe and Mail, has the highest percentage and S. Korea newspaper, Hankyoreh, is the lowest percentage. In terms of Joy, S. Korea newspaper showed the highest percentage and N. Korea newspaper showed the lowest percentage. In

Table 9 Emotion result of S. Korea and Park

News	Anger (%)	Disgust (%)	Fear (%)	Joy (%)	Sadness (%)	Surprise (%)
1	17.62	9.69	28.16	18.22	13.34	12.93
2	19.15	7.52	25.07	16.56	20.56	11.12
3	20.14	8.60	26.75	17.65	15.52	11.31
4	19.18	10.49	31.14	17.70	11.96	9.5
5	20.49	7.59	28.71	18.08	13.47	11.63
6	26.48	15.21	27.09	7.04	16.52	7.62
7	19.51	7.31	28.78	11.38	19.34	13.65
8	21.55	10.26	30.64	12.46	14.51	10.55

Chosun = 1, Hankyoregh = 2, Korea times = 3, Globle and mail = 4, People china = 5, Uriminzokkiri = 6, Thanhnien = 7, Newyork times = 8

Table 10 K-means results of S. Korea and Park

News	K-value 2	K-value 3
Chosun	1	1
Hankyoreh	1	1
Korea Times	1	1
Globe and Mail	2	3
People China	2	3
Uriminzokkiri	1	1
Thanhnien	1	2
Newyork Times	1	1

terms of words including Sadness, Hankyoreh media has the highest figures and the Canada media has the lowest figures. In terms of Surprise, the Thanhnien has the highest number of results. Based on this result, we can see it as a radial chart difference.

For objective analysis, we clustered each news through K-means technique mentioned above. As shown in Table 10, when K is 2, Canadian and Chinese newspaper are grouped together, and when K is 3, Vietnamese news media are different.

Finally, we find that each news report presents the same issue, but the results are different depending on the state, nature and culture of the country.

5 Conclusion

We experimented to identify the framing theory of each country. Thus, eight newspapers were selected from six countries. Through emotional analysis, we analyzed the news articles related to Kim Jung-Un and Park Geun-Hye in each country, and the system is visualized how we think about each person in each country. Though the K-means algorithm, visualization is proved.

The first of all, in relation to Kim Jong-un, South Korea, Vietnam, Canada and US news showed a similar pattern when K-value was 2. However, it can be seen that the news of China and North Korea are different kinds of news. In other words, it can be thought that the news in North Korea and China were framing the same about Kim Jong-un at that time. If the K-value is 3, we could have different values only in the Vietnam press. This means writing an article similar to the South Korean, American, and Canadian press, but it can be said that the report has a slight difference. As a whole, we can think that China and North Korea media produce the same type of news related to Kim Jong-un.

The second of all, it is news related to former President Park Geun-hye. The K-value of 2 would have resulted in Canada and the Chinese press producing similar news. Although it is ideologically different, the news in North Korea can be said to produce news in a similar form to the Republic of Korea. At that time, the government

of Park Geun-hye had been giving a hard - line policy to North Korea and the crisis of lame duck was coming in Korea. For this reason, North Korea has produced similar news related to Park Geun-hye. Chosun is South Korea's representative conservative newspaper, and Hankyoreh is South Korea's representative leftist newspaper. However, we can confirm that both newspapers produce news with a similar position in relation to Park.

In conclusion, we made an effort to identify differences in news production in each country. Especially, we wanted to analyze the news of divided South Korea and North Korea divided by the difference of thought and selected the people who represent each country. We analyzed the news through text mining, and based on it we were able to identify the differences in news production.

References

1. Andsager, J.L., Powers, A.: Social or economic concerns: how news and women's magazines framed breast cancer in the 1990s. Journal. Mass Commun. Q. **76**, 531–550 (1999)
2. Durham, F.D.: Framing the state in globalization: the financial times' coverage of the 1997 Thai currency crisis. Crit. Stud. Mass Commun. **24**(1), 57–76 (2007)
3. Edelman, M.J.: Contestable categories and public opinion. Polit. Commun. **10**, 231–242 (1993)
4. Entman, R.M.: Framing U.S. coverage of international news: contrasts in narratives of the KAL and Iran air incidents. J. Commun. **41**, 6–27 (1991)
5. Entman, R.M.: Framing bias: media in the distribution of power. J. Commun. **57**, 163–173 (2007)
6. Gamson, W.: Talking Politics. Cambridge University Press, New York (1992)
7. Gitlin, T.: The Whole World is Watching: Mass Media in the Making and Unmaking of the New Left. University of California Press, Berkeley (1980)
8. Goffman, E.: Frame Analysis. Free Press, New York (1974)
9. Goss, B.M.: All our kids get better jobs tomorrow: the North American free trade agreement in the New York Times. Journal. Commun. Monogr. 3–47 (2001)
10. Hall, A.: The mass media, cultural identity and perceptions of national character: an analysis of frames in US and Canadian coverage of audiovisual materials in the GATT. Gazette **62**, 231–249 (2000)
11. Jha, S.: Exploring Internet influence on the coverage of social protest: content analysis comparing protest coverage in 1967 and 1999. Journal. Mass Commun. Q. **84**(1), 40–57 (2007)
12. Kim, K.K.: It's all about trade: United States press coverage of cigarette exports talks and policy. Mass Commun. Soc. **6**(1), 75–97 (2003)
13. Solomon, W.S.: News frames and media packages: covering El Salvador. Crit. Stud. Mass Commun. **9**, 56–74 (1992)
14. Entman, R.M.: Framing: toward clarification of a fractured paradigm. J. Commun. **43**, 51–58 (1993)
15. Tankard, J., Hendrickson, L., Silberman., J., Bliss, K., Ghanem, S.: Media frames: approaches to conceptualization and measurement. In: Paper Presented at the Meeting of the Annual Convention of the AEJMC, Boston, MA (1991)
16. Scheufele, D.A.: Framing as a theory of media effects. J. Commun. **49**, 103–122 (1999)
17. Kim, K.K.: Selling smoke to Asia: historical analysis of conflicting U.S. polices on cigarette advertising and promotions. Asian J. Commun. **12**(2), 120–140 (2002)
18. Kim, S.T.: Making a difference: U.S. press coverage of the Kwangju and Tiananmen prodemocracy movements. Journal. Mass Commun. Q. **77**, 22–36 (2000)

19. Maslog, C.C., Lee, S.T., Kim, H.S.: Framing analysis of a conflict: How newspapers in five Asian countries covered the Iraq war. Asian J. Commun. **16**(1), 19–39 (2006)
20. Tian, Y., Stewart, C.M.: Framing the SARS crisis: a computer-assisted text analysis of CNN & BBC online news reports of SARS. Asian J. Commun. **15**(3), 289–301 (2005)
21. Chang, K.: Auto trade policy and the press: auto elite as a source of the media agenda. Journal. Mass Commun. Q. **76**, 312–324 (1999)
22. Manning, C.D., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S.J., McClosky, D.: The Stanford CoreNLP natural language processing toolkit. In: Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations, June 2014, pp. 55–60
23. Mohammad, S.M., Turney, P.D.: Crowdsourcing a word-emotion association Lexicon (2013)
24. Vattani, A.: K-means requires exponentially many iterations even in the plan
25. http://www.internetworldstats.com/stats.htm
26. http://english.chosun.com
27. http://www.koreatimes.co.kr/www2/index.asp
28. http://english.hani.co.kr/
29. http://www.theglobleandmail.com/
30. http://en.people.cn/
31. http://www.uriminzokkiri.com/
32. http://www.thanhniennews.com/
33. http://www.nytimes.com/
34. Android Development Tools. http://www.mkyong.com/java/jsoup-html-parser-helloworld-examples
35. https://developer.nytimes.com/
36. Choi, J., Han, Y., Kim, Y.: A research for finding relationship between mass media and social media based on agenda setting theory. In: Software Engineering Research, Management and Applications, May 2016, pp. 103–113
37. Lasecki, W.S., Rello, L., Bigham, J.P.: Measuring text simplification with the crowd (2015)
38. Galiotou, E., Karanikolas, N., Tsoulloftas, C.: On the effect of stemming algorithms on extractive summarization: a case study (2013)
39. Korenius, T., Laurikkala, J., Jarvelin, K., Juhola, M.: Stemming and lemmatization in the clustering of finnish text documents
40. Ekman, P.: An argument for basic emotions. Cogn. Emot. **6**(3), 169–200 (1992)
41. Mohammand, S.M., Turney, P.D.: Emotions evoked by common words and phrases: using mechanical Turk to create an emotion lexicon

A Travel Decision Support Algorithm: Landmark Activity Extraction from Japanese Travel Comments

Siya Bao, Masao Yanagisawa and Nozomu Togawa

Abstract To help people smoothly and efficiently make travel decisions, we utilize the advantages of travel comments posted by thousands of other travelers. In this paper, we analyze the feasibility of exploring landmark activity queries and representative examples from Japanese travel comments. Contributions in this paper include a framework for extracting activity concerned keywords and queries, quantifying the relationship between landmark activities and comment contents. An evaluation of activity-example extraction is conducted in two case studies through 18,939 travel comments.

Keywords Travel comment · Activity keyword · Query construction · Activity-example

1 Introduction

Travel decisions are grounded in the real experiences of users' travelogues. Thus, investigations on travelogues potentially leads to more reliable travel decisions than simply browsing web pages [13]. Many existing studies utilized travelogues to predict travel activities/purposes using open-domain data such as Twitter and Facebook. These studies includes activity classifications [2, 9], location extractions [5] and daily activity pattern predictions [3, 6]. Cui et al. [2] predicted general 4 types of activities for each landmark using Twitter contents. Lian et al. [9] recognized users' activity types based on check-in and transition histories. Hoang et al. [5] identified location-related information in social media contents. Cranshaw et al. [3] focused on

S. Bao (✉) · M. Yanagisawa · N. Togawa
Department of Computer Science and Communications Engineering, Waseda University, Tokyo, Japan
e-mail: siya.bao@togawa.cs.waseda.ac.jp

M. Yanagisawa
e-mail: myanagi@waseda.jp

N. Togawa
e-mail: togawa@togawa.cs.waseda.ac.jp

R. Lee (ed.), *Computer and Information Science*, Studies in Computational Intelligence 849, https://doi.org/10.1007/978-3-030-25213-7_8

mapping daily activity area patterns in urban cities and Hasan et al. [6] used geo-location data to identify activity categories and then predicted weekly activity patterns of individual users. However, to fully explore the value of large-scale travelogues in travel decision making, there still are issues remaining unsolved.

(1) Dissemination of information: These travelogue data were collected directly from SNS (Social Networking Service) which contains a large amount of irrelevant information including promotions and advertisements [1, 11]. This disadvantage may result in the time-consuming data cleaning process. Thus, it is important and efficient to use reliable data such as travel comments on trustworthy travel sites such as TripAdvisor [4].

(2) Limited information-searching capability: Compared with common information searching, it may be too difficult for a user who is not familiar with the destination city or landmark to raise a question. For example, when visiting an art museum, people often intend to ask a general question such as *What can I see?* or *What is on the exhibition now?* rather than a specific question such as *Is that painting painted in 1800?*. The input question query such as *Why do people like bass low frequencies on music* in [15] is not available in the travel searching case. Thus, how to automatically generate the query is important.

(3) Ambitious description: After inputting the question query, we should search for information related. Short messages such as tags or event phrases may be helpful for describing the landmarks but still could be less meaningful sometimes. For example, most of the tags for *Sensoji* on Twitter are geo-tags such as *#Tokyo* and *#ExploreTokyo* which is ambitious and does not include any details of activities. Thus, illustrating a series of specific examples, such as *The 80's painting exhibition is great!*, will better help users make travel decisions.

(4) Lack of Japanese analysis: Furthermore, most of the query searching problems are implemented in English while the differences among languages are not considered yet. Thus, to fill in the blank in Japanese information searching. We focus on the investigation of travel comments written by Japanese users.

In this paper, to deal with those issues mentioned above, we propose an algorithm to construct activity queries and extract meaningful examples as detailed descriptions. Our contributions are highlighted as follows:

- Extract activity keywords based on three types of frequency factors (Sect. 3.3).
- Construct activity queries with considerations on Japanese linguistic features (Sect. 3.4).
- Develop the activity score to select representative activity examples (Sect. 3.4).
- Two case studies applying the algorithm with 18,939 comments on Jaran (a Japanese travel website) [7] (Sect. 4).

2 Data Collection

The goal of our proposed algorithm is to extract interesting activities for a landmark based on a large corpus of travelogues. In other words, we ensure our algorithm so as to match the domain nature of travelogues and problem domain as much as possible. Instead of using travelogues such as tweets of which contents are not limited to travel experiences but also other reports such as food, sports, and health, we set our data domain as a corpus of travel comments obtained from the professional travel website. The advantages of using travel comments are as follows:

1. Using travel comments is efficient and convenient as the traditional data-cleaning process conducted by crowdsources is no longer needed. This is because non-experiential information such as advertisements in Tweets rarely exists in travel comments on big travel websites [4].
2. A travel comment usually is longer than a tweet which has the limitation of 140 characters. The median length of comments that we collected is over 40 Japanese characters (see Table 1). It means that a travel comment usually contains around 2–3 completed sentences in Japanese. This advantage ensures the informativeness of each comment.

To sum up, due to these two advantages, using travel comments is convenient and should guarantee the quality of our experimental datasets.

In this paper, we set Tokyo as the target city, and Jaran [7], which is domestic travel website, as the target website in Japan. We focused on 5 famous public landmarks in Tokyo on Jaran, which are

- 三鷹の森ジブリ美術館 (Ghibli Museum)
- 浅草寺 (Senso Temple)
- 上野動物園 (Ueno Zoo)

Table 1 The statistic of travel comments

Landmark (Japanese)	三鷹の森ジブリ美術館	浅草寺	上野動物園	国営昭和記念公園	明治神宮
Landmark (English)	Ghibli Museum	Senso Temple	Ueno Zoo	Showakinen Park	Meiji Jingu
Category	Museum	Temple	Zoo	Park	Shrine
Number of comments	1936	7101	4170	2531	3201
Max length of comments[a]	334	391	517	382	391
Median length of comments[a]	41	40	42	41	41
Min length of comments[a]	5	4	4	3	3

[a]Counted in Japanese characters

- 国営昭和記念公園 (Showakinen Koen)
- 明治神宮 (Meiji Jingu)

For each landmark, we collected all travel comments of it through its full timeline. Table 1 summarizes the data statistic of travel comments collected. All travel comments were written before November, 2018[1] and were collected through a data-collection program in R. Comments written in English (less than 1%) and comments only contained photos were not included. As a result, 18,939 travel comments were accumulated in Japanese.

3 Activity-Example Extraction

3.1 Problem Definition

Given a landmark, the goal is to explore potential activity descriptions that may benefit users in travel decision making. The inputs contain a landmark name and a large corpus of travel comments in Japanese. The outputs are the most likely activity query and top-5 examples in the form of an activity-example table. In summary, our activity-example extraction algorithm for a landmark is described as follows:

Step A1: Phrase all comments in the entire comment set.
Step A2: Obtain the keyword using the algorithm described in Sect. 3.3.
Step A3: Construct the query with the keyword using the algorithm described in Sect. 3.4.
Step A4: Extract top-5 examples containing the query using the algorithm described in Sect. 3.5.

3.2 Linguistic Features

Japanese is quite different from English. For the sake of clear understanding, we introduce three basic structures in Japanese [10]. Unlike English, there is no space between words in Japanese. For this reason, in order to better understand the structures of Japanese, we separate each important elements for each sentence in the following three examples.

1. 金さん は　男性です。(Kin is a male.)
2. 金さん が 走る。(Kin runs.)
3. 金さん は　優しいです。(Kin is thoughtful.)

Those structures can be generally considered as *S (Subject) + ST (Statement)* as Japanese is a topic-prominent language [12]. In example (1), *ST* is *male* which is a *N*

[1]The date is users' visiting date, not the written date.

Table 2 Example of element phrasing

Example 1	ジブリの作品が好き (Like the works by Ghibli)				
Element	ジブリ	の	作品	が	好き
POS	Noun	AUX	Noun	AUX	Adjective verb

Example 2	ジブリの世界観に浸れる (Immersed into the world of Ghibli)				
Element	ジブリ	の	世界観	に	浸れる
POS	Noun	AUX	Noun	AUX	Verb

noun. In example (2), *ST* is *run* which is a *V* verb. In example (3), *ST* is *thoughtful* which is an *A* adjective.

Moreover, there are usually *AUX* auxiliary words between segments in Japanese. In example (1), は is the *AUX*, and it connects the parts 金さん *(Kin)* and 男性 *(male)* and is similar to the meaning of *is* in English. In example (2), が is the *AUX*, and it connects the parts 金さん *(Kin)* and 走る *(run)* which emphasizes the subject 金さん *(Kin)*.

Thus, our goal is to construct a query with a structure of *S (Subject) + ST (Statement)* and we consider *S* is a noun.

For each comment, we extract the dependency relationships and part of speech tags (POS) through RMeCab [14]. Table 2 shows two examples of phrased travel comments in Japanese. Note that, we do not consider the differences of tenses in the phrasing process. Also, POS including prefix and postfix are not considered either.

3.3 Keyword Finding

Based on the discussion in Sect. 3.2 above, to construct a query, we first identify the noun N for the subject S. N should be representative, in other words, N is the keyword for the landmark.

To find the keyword, we consider words frequently associated with a landmark (TF), while not with so many other landmarks (IDF). To quantify the relevance between a noun and a landmark, we define a "document" as one travel comment in the comment set I_l which contains all comments for the landmark l. Then, we evaluate each noun $n \in I_l$ via TF and IDF.

Specifically, we use three types of frequencies: noun frequency (nf), comment frequency (cf), and landmark frequency (lf). Given a noun n and a landmark l, $nf(I_l, n)$ is the number of appearances that have n in I_l. It is reasonable that a noun is likely to be one of the key nouns as more comments of the landmark l have the noun. However, some comments have the same nouns in the same comment causing overestimation of nf. Likewise, $cf(I_l, n)$ is the number of comments that include n

Table 3 Example of the keyword ranking for Ueno Zoo

Noun	Rank
パンダ (Panda)	1st
動物園 (Zoo)	2nd
子供 (Child)	3nd

in I_l. *cf* avoids overestimating *nf*. However, we need to filter common nouns like 人(*people*) and 私（*I*), which also have high *nf* and *cf*. To deal with the problem, given a noun *n* and *L*, which is a set *l* of all landmarks, let $lf(I_L, n)$ be the number of comments having *n* in the entire comment set I_L.

Using these three types of frequencies, we define a keyword noun score *KNS* of a noun *n* as:

$$KNS(n) \propto \frac{nf(I_l, n) \times cf(I_l, n)}{lf(I_L, n)} \tag{1}$$

We consider a noun *n* as the keyword *key* if $KNS(n)$ is the largest for all $n \in I_l$. Table 3 shows an example list of keyword ranking for Ueno Zoo.

In addition, geo-specific nouns which represent a landmark's spatial locations such as *Shanghai* and *China* may frequently appear, but those words are noisy terms as it has fewer associations with activities. In an instance, *Shanghai* often comes with a comment *It is garden/park/building in Shanghai*. To quantify the keyword extraction, if a keyword *key* is a geo-specific noun, then we use the noun with the second high $KNS(n)$ value.

In summary, our keyword find algorithm for a landmark *l* is described as follows:

Step K1: For a landmark *l*, extract the keyword *key* of it as follows:

(1-1) Calculate *nf*, *cf* and *lf* values for a noun *n*.
(1-2) Calculate $KNS(n)$.

Step K2: Select the non-geo-specific noun with the highest $KNS(n)$ value as the keyword *key*.

3.4 Query Construction

After extracting the keyword *key*, we complete the query with the part of statement *ST* through:

1. Clustering: We search for 3-grams phrases using *key* obtained in Sect. 3.3 among *Noun, Verb, Adjective, Adjective verb, Adverb*. Users may express their personal experiences in different ways when writing the comments, and this results in

identifying queries with substantially the same meanings. To handle this problem, we cluster similar 3-grams phrases based on the cosine similarity. If the cosine similarity between two 3-grams phrases is over the threshold d_{th} (we set $d_{th} = 0.67$ in the experiment), then we merge the phrases. The cluster name is the most repeated elements in the cluster.

2. Query Representation: We extract the most hit element phrase from the cluster. Then, we predict the AUX between any two elements in the element phrase. Given two elements e_1, 作品 *(Wokr, Noun)* and e_2, 好き *(Like, Adjective verb)* (see query representation in Fig. 1), let A be a set of all *aux*'s in the I_L, and for $aux \in A$, we define two probability factors:

   - The first probability factor is generic probability $gp(I_L, aux, e_1, e_2)$, it is the probability of *aux* connects *Noun* and *Adjective verb* in the entire comment set I_L.
   - The second probability factor is specific probability $sp(I_l, aux, e_1, e_2)$, it is the probability that *aux* connects 作品 *(Wokr)* and 好き *(Like)* in the comment set of the landmark I_l.

With these two types of probability factors, we define a connection probability $cp(aux, e_1, e_2)$ predicting the relationship between *aux* and two elements e_1, e_2 as follows:

$$cp(aux, e_1, e_2) \propto gp(I_L, aux, e_1, e_2) \times sp(I_l, aux, e_1, e_2) \tag{2}$$

We consider that the *aux* with the highest $cp(aux, e_1, e_2)$ value connects e_1 and e_2. By repeating connecting all elements in the element phrase, we finally obtain a complete query q.

3.5 Example Searching

The next step is to find representative comments match the query q. We define a scoring mechanism called activity score AS to select example comments. We select comments containing or partly containing the query q and score them by AS.

To score each comment, we use three factors:

1. *Comparison factor*: It is widely used in decision making [8]. This factor is computed as a binary variable. It is 1 when the query q appears in the comment and otherwise it is 0.
2. *Coverage factor*: It is defined as the percentage of a query q contained in a comment and it is normalized between 0 and 1 (see Fig. 2).
3. *Length factor*: A comment with a longer length is more informative. The factor is also normalized between 0 to 1 compared with all comments in I_l. The longest comment is 1 and the shortest one is 0.

We define the AS through a weighed linear relationship:

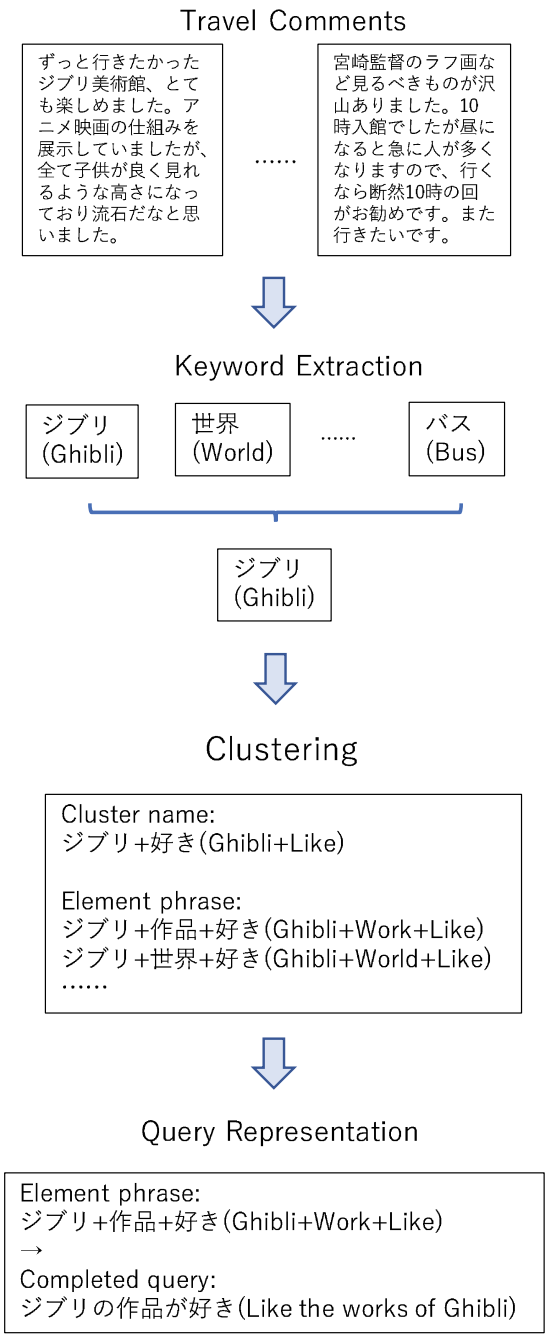


Fig. 1 Query construction

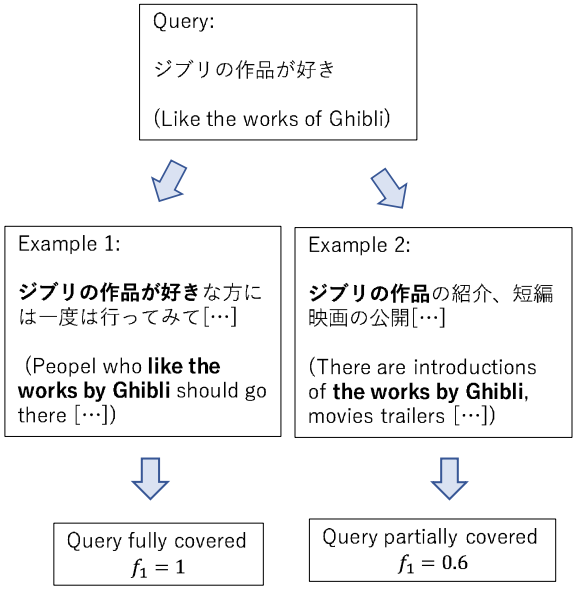


Fig. 2 Coverage factor

$$AS = w_1 f_1 + w_2 f_2 + w_3 f_3 \tag{3}$$

when w_i is the weight for ith factor. f_1, f_2, f_3 represents comparison, coverage and length factors, respectively. At last, top-5 comments are picked as the examples for detailed descriptions of the landmark activity.

4 Case Studies

4.1 Evaluation Criteria

In this section, we evaluate two case studies by deriving meaningful activity queries and example comments. We evaluate the quality of experimental results in two terms:

- Whether the constructed queries are meaningful
- Whether the extracted examples are proper.

4.2 Data

Our goal is to process a full dataset of travel comments, but we first implement the small-scale test to confirm the effectiveness of our proposed algorithm.

We use the 18,939 collected in Sect. 2 as the entire comment set I_L.

For the first case study, we have a subset of all Japanese travel comments for the Ghibli Museum with 1936 comments during the period of May 2007 to October 2018.[2] For the second case study, we have a subset of all Japanese travel comments for the Ueno Zoo with 4170 comments during the period of April 2006 to October 2018.[3]

4.3 Procedures

We invited 5 volunteers as evaluators. 3 of them are in computer science and engineering fields. 2 of them are English speakers, and the other 3 volunteers are Japanese and English speakers. All of them are not authors of this paper.

For each study case, we derived one query and corresponding examples with top-5 *AS* values. The query and examples were originally in Japanese and translated into English lately. Volunteers were asked to first rate to what extent they understood the queries with a scale from 1 (do not understand)–5 (fully understand), and then, they were asked to rate the relevance of the top-5 examples, from 1 (least relevant)–5 (most relevant). Based on the ground-truth, we calculated normalize Discounted Cumulative Gain ($nDCG$) to evaluate the quality of the obtained examples.

4.3.1 Evaluations on Query Construction

For the evaluation of query quality, we compared our proposed algorithm with two baseline algorithms with disabling parts of our algorithm as follows:

- Keyword query: It has the most hit noun as a query based on steps from Sect. 3.3.
- Incomplete query: It constructs a query based on steps from Sects. 3.3 to 3.4 without predicting the AUX's.

Table 4 shows the results of query evaluation. Generally speaking, it is observed that our algorithm outperforms the other two algorithms in both case studies. Our algorithm is at least 10% higher than the other two baseline algorithms. It is interesting to mention that in case 2, there is no AUX between 一番 *(most)* and 人気 *(popular)* by our proposed algorithm due to the common usage in Japanese.

- Ours versus Keyword query: Since the word *Panda* is more comprehensive compared to the word *Ghibli*, which is a Japanese word. Thus, Keyword query obtained a higher score in case 2 than that in case 1. On the other hand, it indicates that our algorithm is useful to help users have a more comprehensive image of the landmark activity dealing with descriptions in the foreign languages that they are not familiar with.

[2] https://www.jalan.net/kankou/spt_13204cc3302011245/.

[3] https://www.jalan.net/kankou/spt_13106cc3310040182/.

Table 4 Results on query construction

Case 1: Ghibli Museum		
Algorithm	Avg. score	Query
Ours	4.2	ジブリの作品が好き） (Like the works by Ghibli)
Keyword query	2.6	ジブリ (Ghibli)
Incomplete query	3.2	ジブリ+作品+好き (Ghibli+Work+Like)

Case 2: Ueno Zoo		
Algorithm	Avg. score	Query
Ours	4.4	パンダが一番人気 (Panda is the most popular)
Keyword query	3.4	パンダ (Panda)
Incomplete query	4.0	パンダ+一番+人気 (Panda+Most+Popular)

- Ours versus Incomplete query: Although Incomplete query provides more information, our algorithm has a completed query rather than a bag of words. Thus, it is reasonable that our algorithm is easier to understand and obtains higher scores in both case studies.

4.3.2 Evaluations on nDCG

To extract the top-5 examples, we evaluated each comment using AS. We set the following weights for the three factors in AS with $w_1 = 0.25$, $w_2 = 0.5$ and $w_3 = 0.25$, which has the best performance according to the prior experiments.

Then, DCG is computed firstly as follows:

$$DCG_p = r_1 + \sum_{i=2}^{p} \frac{r_i}{log(i)} \tag{4}$$

where r_i is the score of ith example, and DCG_p is the accumulated score at ith example.

As DCG depends on the size of the extracted examples, the more examples are extracted, the larger DCG becomes. In other words, it is not explicitly to figure out

Table 5 $nDCG_p$ of our proposed algorithms in two case studies

	Case 1: Ghibli Museum	Case 2: Ueno Zoo
Avg. $nDCG_1$	1.000	1.000
Avg. $nDCG_2$	0.900	1.000
Avg. $nDCG_3$	0.920	0.952
Avg. $nDCG_4$	0.966	0.991
Avg. $nDCG_5$	0.969	0.992

whether the extracted examples are qualified or not. Thus, to further normalize the results we compute $nDCG$ either as follows.

$$nDCG_p = \frac{DCG_p}{IDCG_p} \tag{5}$$

where $IDCG_p$ is the ideal rank.

The results are listed in Table 5. Our algorithm achieves high $nDCG$ values in both case studies at any ranks. It indicates that our extracted examples are quite relevant to the landmark activity descriptions and can help users better understand what is interesting to do and see when visiting the landmark. Generally speaking, according to the experimental results, our algorithm could benefit users in the better travel decision-making process.

4.3.3 Extracted Examples

The top-5 examples extracted for two case studies are listed in Tables 6 and 7. The queries contained in the examples are bold.

In case 1, 1st example to 4th example contains the complete query (ジブリの作品が好き *(Like the works by Ghibli)*), while 5th example contains a part of the query (ジブリの作品 *(The works by Ghibli)*). For better understanding, English translations of first to fifth examples of Ghibli Museum are listed as follows:

- *People who **like the works by Ghibli** should go there. Reservation is needed. It does not only include Totoro but also includes other works of Ghibli[...].*
- *Pretty **like the works by Ghibli** a long time ago and I often watch the movies. It feels like getting in the world of Ghibli [...].*
- *[...] I like **the works by Ghibli**, I have been there for several times. Whenever you go to the museum, It feels like going back to your childhood time [...] .*
- *[...] I like to watch to **the works by Ghibli** in my spare time, movies trailers. If you go there, it feels like getting in the world of Ghibli [...] .*
- *[...] There are introductions of **the works by Ghibli**, movies trailers, models of the cat bus, Castle Laputa in the Sky [...] .*

Table 6 Example comments for Ghibli-museum

Query	Examples	*AS*
ジブリの作品が好き (Like the works by Ghibli)	ジブリの作品が好きな方には一度は行ってみて [...] トトロだけではなく他作品も含めて [...]	0.782
	昔からジブリの作品が好きでよく見ています。ここはジブリの世界に入り込んだ [...]	0.774
	ジブリの作品が好きなので、リピートしています。何度行っても、童心に帰れて [...]	0.774
	[...] ジブリの作品が好きで暇さえあれば見ています。ジブリ美術館に行くとジブリの世界 [...]	0.772
	[...] ジブリの作品の紹介、短編映画の公開、ネコバス、天空の城ラピュタの再現 [...]	0.717

Table 7 Example comments for Ueno Zoo

Query	Examples	*AS*
パンダが一番人気 (Panda is the most popular)	[...] パンダが一番人気であることが分かります。食事中のパンダは [...]	0.810
	[...] パンダが一番人気でとても可愛いくて癒されます。アリクイなど他にも [...]	0.779
	[...] パンダが一番人気だけど、珍しい変テコな動物もいっぱいいてビックリします [...]	0.774
	[...] パンダが一番人気だと思います。入り口入って右側にパンダのエリアがある [...]	0.774
	[...] パンダが一番人気で多くの人が行列をして見ていました。パンダがえさを食べてる姿 [...]	0.771

In case 2, all examples contain the complete query (パンダが一番人気 *(Panda is the most popular)*). For better understanding, English translations of first to fifth examples of Ghibli Museum are listed as follows:

- *[...]* ***Pandas are the most popular****. It is great to have a chance to see panda slowly eating[...].*
- *[...]* ***Pandas are the most popular*** *and are so cute that heal the soul. There are other animals such as anteater [...].*
- *[...]* ***Pandas are the most popular****, also it is so surprising that there so many other rare animals [...].*
- *[...] I think* ***Pandas are the most popular****. The panda area is on the right side of the zoo entrance [...].*
- *[...] Because* ***Pandas are the most popular****, I saw them after waiting a lot in lines. Pandas were eating their food [...].*

5 Conclusion

In this paper, we concentrate on answering the question: Can we accurately and directly tell users what to do and see during their visits via previous travel comments?

Based on the experimental results, with a relatively small scale of travel comments, it still allows us to explore rich landmark activity information.

For future works, we are continuing to deepen our works and intend our scope to other language-specific travel comments such as English and Chinese. Meanwhile, the workloads of the applications will be evaluated.

Acknowledgements This paper was supported in part by Grant-in-Aid for Scientific Research (No. 19H04080).

References

1. Boyd, D., Golder, S., Lotan, G.: Tweet, tweet, retweet: conversational aspects of retweeting on Twitter. In: Proceedings of International Conference on HICSS (2010)
2. Cui, Y., Meng, C., He, Q., Gao, J.: Forecasting current and next trip purpose with social media data and Google Place. Transp. Res. Part C Emerg. Technol. **97**, 159–174 (2018)
3. Cranshaw, J., Schwartz, R., Hong, J.I., Sadeh, N.: The livehoods project: utilizing social media to understand the dynamics of a city. In: Proceedings of International Conference AAAI on WSM (2012)
4. Filieri, R., Alguezaui, S., McLeay, F.: Why do travelers trust TripAdvisor? Antecedents of trust towards consumer-generated media and its influence on recommendation adoption and word of mouth. Tour. Manag. **51**, 174–185 (2015)
5. Hoang, T.B.N., Mothea, J.: Location extraction from tweets. Inf. Process. Manag. **54**, 129–144 (2018)
6. Hasan, S., Ukkusuri, S.V.: Urban activity pattern classification using topic models from online geo-location data. Transp. Res. Part C Emerg. Technol. **44**, 363–381 (2014)

7. Jaran: https://www.jaran.jp
8. Jindal, N., Liu, B.: Mining comparative sentences and relations. In: Proceedings of International Conference on AAAI (2006)
9. Lian, D., Xie, X.: Collaborative activity recognition via check-in history. In: Proceedings of International Conference on LBSN (2011)
10. Makino, A., TSutsui, M.: A Dictionary of Basic Japanese Grammar. The Japan Times, Tokyo (1989)
11. Naaman, M., Boase, J., Lai, C.H.: Is it really about me?: message content in social awareness streams. In: Proceedings of International Conference on CSCW (2010)
12. Ramchand, G., Resis, C.: The Oxford Handbook of Linguistic Interfaces. Oxford Handbooks in Linguistics, pp. 358–362. Oxford University Press, Oxford (2007)
13. Rashidi, T.H., Abbasi, A., Maghrebi, M., Hasan, S., Waller, T.S.: Exploring the capacity of social media data for modelling travel behaviour: opportunities and challenges. Transp. Res. Part C Emerg. Technol. **75**, 197–211 (2017)
14. RMeCab: https://code.google.com/p/mecab/downloads/list
15. Shen, H., Liu, G., Wang, H., Vithlani, N.: SocialQ&A: an online social network based question and answer system. IEEE Trans. Big Data **3**, 91–106 (2017)

Design of Intelligent Artificial Agents: Its Application in Joint Attention Task for Children with Autism

Vishav Jyoti and Uttama Lahiri

Abstract Recently computer-based applications for skill training is gaining popularity. An important ingredient is a mediator. Thus, the design of realistic artificial agents (agent) enacting the role of the mediator is important. The agent needs to look like a real human being demonstrating naturalistic behavior. The agent can possess complexion, hair color, etc. depicting the user's nationality. Also, the agents are empowered with intelligence such as in agent-mediated joint attention (JA) task. The JA skill refers to one's ability to attend to an object in the shared space with a social partner. However, currently existing computer-based agent-mediated JA tasks either use agents that are not human-like or that show only the face, stealing realism. In our present research, we have designed agents with Indian look and capable of behaving naturally. Also, the intelligent engine empowers them to intelligently administer JA tasks. Results of a study with six typically developing (TD) children and six with autism spectrum disorder (ASD) are promising.

Keywords Autism · Intelligent agent · Joint attention · Prompting cues

1 Introduction

With the improvement in computational power coupled with rich graphics, computer-based applications are becoming more realistic. For effective interaction, an embodiment in the form of a virtual character might be critical. Thus, the design of a realistic virtual character, considered as an artificial agent (agent *henceforth*) is emerging as one of the upcoming areas in computer-based research. Some of the important traits characterizing a virtual agent are its close resemblance to a real human being [1], such as having a look that represents the same nationality as the user to which the user can easily relate; is capable of demonstrating subtle but naturalistic behavior

V. Jyoti (✉)
Centre for Cognitive and Brain Sciences, IIT Gandhinagar, Gandhinagar 382355, Gujarat, India
e-mail: vishav.jyoti@iitgn.ac.in

U. Lahiri
Department of Electrical Engineering, IIT Gandhinagar, Gandhinagar 382355, Gujarat, India
e-mail: uttamalahiri@iitgn.ac.in

R. Lee (ed.), *Computer and Information Science*, Studies in Computational Intelligence 849, https://doi.org/10.1007/978-3-030-25213-7_9

e.g., blinking, speaking with lip-synching, displaying animations, etc. Agent-human interaction is gaining popularity in different applications. For ex-ample, while administering agent-mediated joint attention (JA) skill training, researchers have used intelligent agents as mediators [2]. The JA refers to the ability of an individual to attend to a certain stimulus (e.g., an object of interest) in the shared space with one's social partner [3]. Specifically, in one of the studies [2], the agent administering a JA task is capable of intelligent reciprocal interaction with the children while offering JA tasks. However, such agents used for computer-based JA tasks have the cartoon-like appearance, though it can perform some basic functions such as head turn and finger pointing. Apart from JA tasks, researchers have also used cartoon characters such as pink dolphins as agents to assist learning of various social skills for children with ASD [4]. There is another study [5] that have used an intelligent agent for administering computer-based JA tasks. Though the agent used in this study is not like a cartoon, yet, the tasks project only the head of the agent that dampens the humanistic appeal thereby adversely affecting the realistic feel to the users. Mostly, these studies have been targeted for individuals with autism spectrum disorder (ASD). This is because, children with ASD are often characterized by deficits in JA skill [3] that can lead to cascading effects in one's language development, crucial for social communication [3] and earning their own livelihood. Designing of platforms that can offer JA-related tasks to this target group is important given the high prevalence of ASD such as 1 in 68 in India [6]. However, as a precursor to developing such a platform, it is essential to design one of the basic components namely an agent who will resemble a 3-D humanoid that is capable to elicit human-like behavior and has the intelligence to autonomously administer a JA task while enacting the role of a mediator. Though it is well established that children with ASD find it difficult to interact with other individuals during social interaction [7], it is essential to train these individuals while being immersed in real-life social scenes featuring humanoid characters. This is because the presence of such characters can lend a touch of reality to the training platform. Given that these children have issues with generalization [8], skill training in presence of realistic virtual humanoid characters might make it easier for them to translate the skills learned from the virtual world to the real world.

In our present work, we have designed a small database of virtual agents. Our agents were designed to look like human beings having Indian nationality. Additionally, the agents were capable of blinking, speaking with lip-synching, displaying animations like turning head, making pointing gestures, etc. Also, we have designed agent-mediated JA tasks in which the agent could intelligently deliver varying JA-related prompting cues using gaze orientation, turning head, lifting hand with a pointed finger towards an object of interest. The agent was made intelligent by using an intelligent prompt protocol engine that empowered the agent to understand whether a user is capable of deciphering a JA-related cue. Subsequently, the agent autonomously take decision on the next JA cue to be delivered, similar to that followed by expert therapists in conventional techniques [9].

The objectives of our study are three-fold namely, to design (i) an agent that looks Indian and appears as a 3-D humanoid capable of demonstrating the subtleties and

animation in a naturalistic manner, (ii) an intelligent prompt protocol engine that can lend intelligence to our agent and (iii) agent-mediated JA tasks.

The rest of the paper is organized as follows: Sect. 2 covers the system design of our research study. Section 3 discusses the materials and methods used in our research work. Section 4 explains the results of our proposed research study and finally, Sect. 5 presents the conclusion and future scope of our research.

2 System Design

Our computer-based JA tasks comprised of (i) Agent, (ii) Virtual Environment (VE) and (iii) Intelligent Prompt Protocol Engine (IPPE).

2.1 Design of Artificial Agent (Agent)

Though there are research studies [2, 5] that have used virtual agents for administering JA tasks, yet none of these have used Indian faces. Our present research was carried out in the Indian context that needed the design of agents with an Indian look. For this, we focused on the characteristic features e.g., color and texture of skin, hair, eyebrows, eyeball, etc. Refining these finer aspects of the agent is important for realistic rendering so that a user can easily relate himself/herself with the agent. These finer aspects were included as an integral part of the major components to building an agent having (i) body (ii) face with morphs and (iii) animation.

2.1.1 Design of Agent's Body

Design of an agent's body is the first step to building an agent. For this, we imported skeleton (see Fig. 1) (comprising of easily modifiable mesh and bones) along with attire (whose attributes can be modified by the designer) from cadnav models (available free of cost) [http://www.cadnav.com/] in .max format. Here, we wanted to use both female and male agents and thus we chose the relevant skeletons from cadnav models. These skeletons were then accessed in 3ds Max [https://www.autodesk.com-/education/freesoftware/3dsmax] environment for developing the desired 3-D agent along with skeleton, clothes, etc. Further, the 3ds Max software was used to modify the attributes of the mesh (of the skeleton) for the rendering of the texture of the agent's body. Using the material editor, we adjusted the skin color (agent's body) so as to give an Indian look to the agent. Our JA task to be presented was designed using the python-based Vizard software (from Worldviz lnc.). Thus, the agent's skeleton was imported in the Vizard platform using the Cal3d exporter software [https://docs.world-viz.com/vizard/latest/cal3d_introduction.htm] following a sequence of steps. First, the Cal3D Skeleton File (.CSF) was imported followed by the Cal3D Mesh

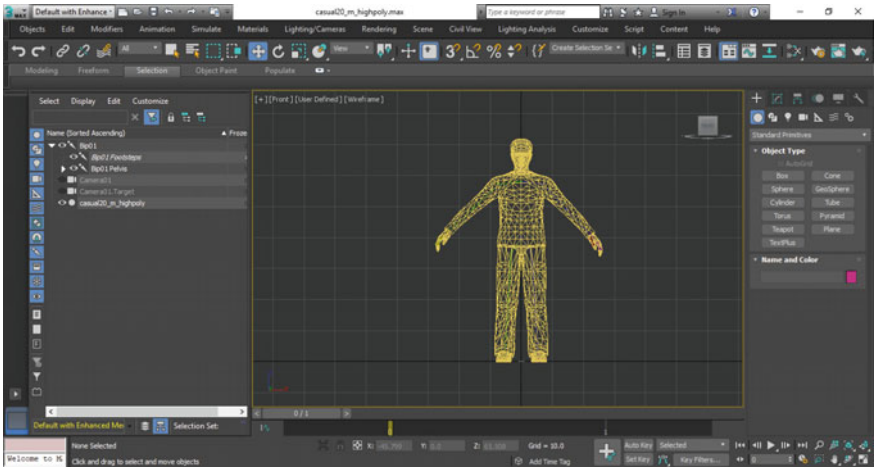


Fig. 1 Agent skeleton in 3ds Max

File (.CMF) that in turn was followed by Cal3D Material File (.CRF). For compatibility with Vizard, an in-house built code was used to generate a file (.cfg format) representing the agent's body in Vizard. The next step was to add a head (with a face) to the agent's body.

2.1.2 Design of Agent's Head

Since our agents were meant to be used by the Indian population, we needed the agent's face to have an Indian look. Also, for the presentation of the faces, we needed 3-D heads. For designing faces, 2-D photographs of front faces from Asian face dataset [http://wiki.cnbc.cmu.edu/Face_Place] were used. For mounting these faces on 3-D heads, we used facegen modeller software [https://facegen.com/-modeller.htm] (see Fig. 2). Using this software, we modified the texture and skin color of the face. Also, we chose an eyeball and hair color to resemble an Indian look. Additionally, we chose the agent's age similar to that of our participant pool by using inbuilt routines of facegen modeller. Finally, we exported our designed heads from facegen modeller in .wrl format to be accessed in the Vizard platform. For our present study, we designed eight heads (4 males and 4 females). Once the head with face was designed, we worked on the agent's emotional expression using face morphing.

Morphing of Agent's Face

For our research study, we wanted to design agents demonstrating neutral faces. The idea was to prevent the use of emotional faces that can offer confound to the JA tasks. Using the morph function available in the PeopleMaker tool (that comes with the

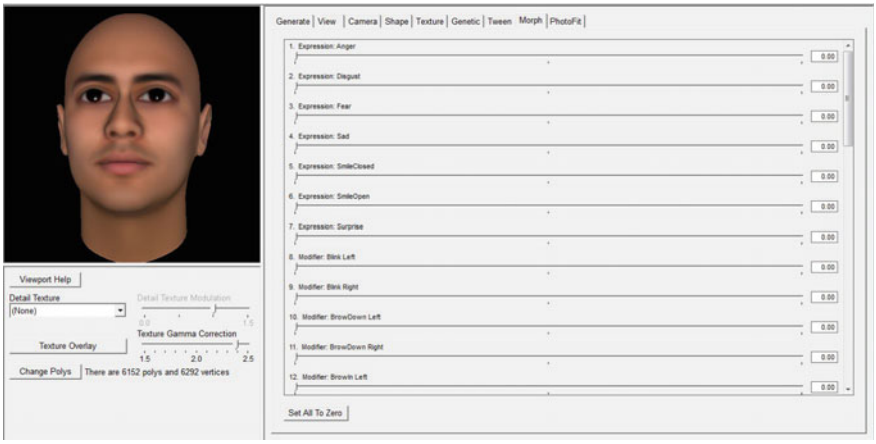


Fig. 2 Agent head creation from 2-D photograph in facegen modeller

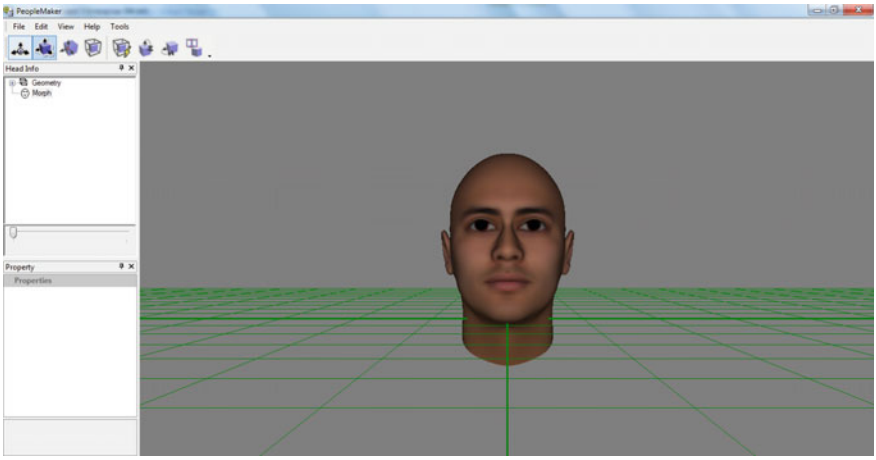


Fig. 3 Agent morph for neutral expression in peoplemaker

Vizard software) we configured the faces to display neutral expression (see Fig. 3). Also, to offer a naturalistic look, we programmed the agents to be capable of blinking randomly with an interval of 1–2 s [10]. Again, we used "speak" morph to enable our agents to speak with lip-synching with pre-programmed audio files.

2.1.3 Stitching of Agent's Head and Body

The agent's body comprises of segments e.g., neck, torso, etc. For stitching the head (see Sect. 2.1.2) with neutral face morph (see Sect. 2.1.2.1) to the body (see

Sect. 2.1.1), we transposed the head to the root bone of the neck of the agent's body for the agent to have a 3-D humanoid embodiment. This was done in the Vizard environment. Figure 4 shows a snapshot of the agents used in our study. Once the agents were designed, we carried out a survey to understand whether the user perceived the agent as was thought by the designer.

Survey Questionnaire Used for the Agents' Faces

For the survey, we used a total of 8 heads comprising of four males (M1–M4) and four females (F1–F4) displaying neutral facial expression. The purpose of the survey was to choose those faces that display neutral expressions and look like Indians to our participants recruited from our neighborhood. The questionnaire was presented on a computer screen with the face on one side and the questions (Q1 and Q2) on the other side of the screen (see Fig. 5). The survey questionnaire (Q1 and Q2) was designed in three languages (English, Gujarati, and Hindi). The question Q1 was asked to ensure whether the respondent felt that the agent demonstrated a specific emotion or neutral expression. The question Q2 was asked to ensure that whether the agent looked Indian to the respondent. Based on the survey, we selected four agents (2 males and 2 females) (see Fig. 4).

Fig. 4 Designed agents

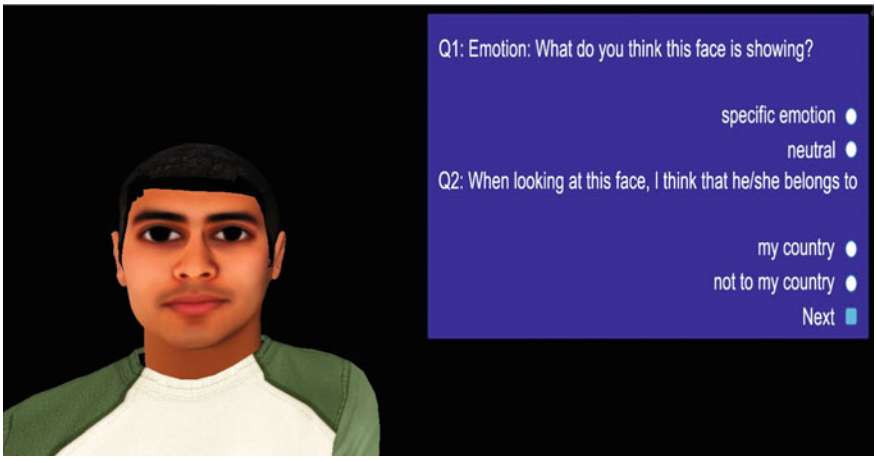


Fig. 5 Interface of survey questionnaire for agent's face

2.1.4 Incorporating Animations

Our agents were designed to demonstrate animations (in Vizard) related to the face, head, and body. As far as the animation of the face was concerned, the agents were programmed to vary the gaze direction. For this, we designed three gaze orientations (θ^0 w.r.t horizontal) i.e., *look straight* ($\theta^0 = 0^0$; see Fig. 6a), *look left* ($\theta^0 = -10^0$; see Fig. 6b) and *look right* ($\theta^0 = +10^0$; see Fig. 6c) where the gaze was programmatically rotated with respect to the center for left and right orientations [11]. Again, for the animation of the head, the agent was programmed to display four types of head turn (α_1^0 is w.r.t horizontal and α_2^0 is w.r.t vertical planes) i.e., *no turn* ($\alpha_1^0 = 0^0$ and $\alpha_2^0 = 0^0$ see Fig. 6d), *left turn* ($\alpha_1^0 = -35^0$ and $\alpha_2^0 = 0^0$; see Fig. 6e), *right turn* ($\alpha_1^0 = +35^0$ and $\alpha_2^0 = 0^0$; see Fig. 6f) and *downside turn* ($\alpha_1^0 = 0^0$ and $\alpha_2^0 = +15^0$; see Fig. 6g). For animations with regard to the body, the agent was programmed to demonstrate finger pointing using index finger while orienting the arms and finger. Thus, we designed three different arm orientations (β_1^0 w.r.t vertical and β_2^0 w.r.t horizontal) i.e., *arms down* (($\beta_1^0 = -75^0$ (left arm), $\beta_1^0 = +75^0$ (right arm)) and $\beta_2^0 = 0^0$) see Fig. 6h, *pointing left* (($\beta_1^0 = -25^0$ and $\beta_2^0 = 0^0$); see Fig. 6i) and *pointing right* (($\beta_1^0 = +25^0$ and $\beta_2^0 = 0^0$); see Fig. 6j). All the angular orientations chosen for incorporating animations in agents (for head and arms) were taken as an initial approximation as per the requirement of our application in which we used our designed agents.

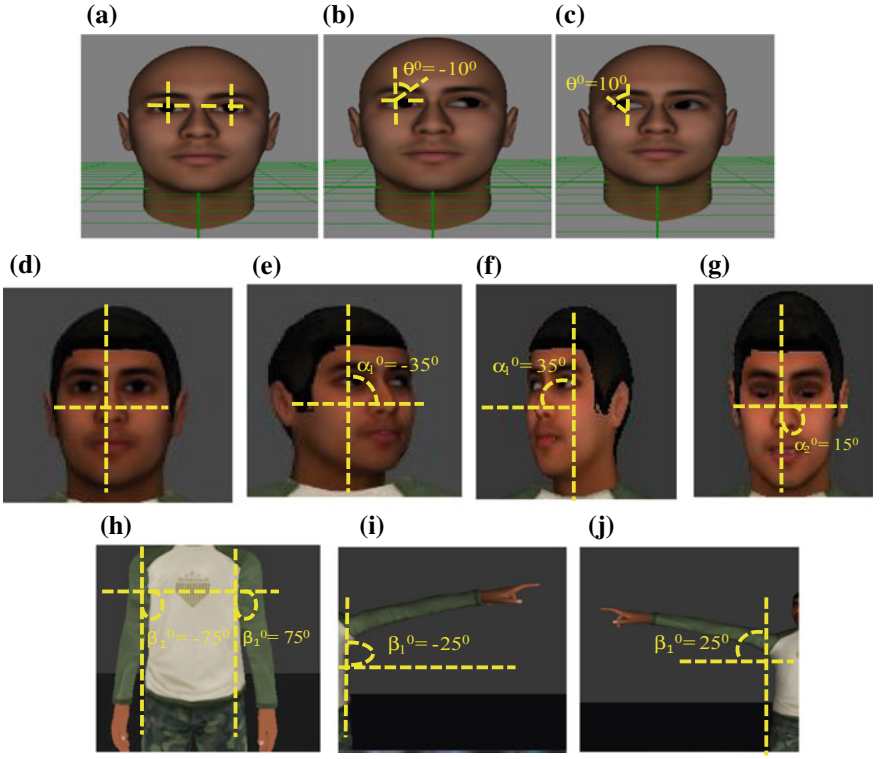


Fig. 6 Agent's animations

2.2 Design of Virtual Environment for Computer-Based Agent-Mediated JA Task

For our computer-based agent-mediated JA task, we needed to design virtual environment (VE). The VE consisted of (i) Virtual room and (ii) Virtual objects (V_{OBJ}).

2.2.1 Virtual Room

For the JA task, we designed a virtual room having 2 shelves hanging from the left and right walls. As can be seen from Fig. 7, an agent stood in the virtual room surrounded by the 2 shelves.

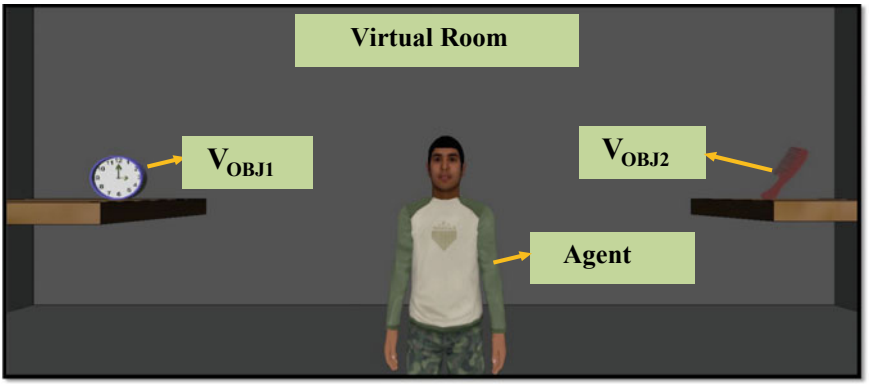


Fig. 7 Virtual room with agent and V_{OBJ}

2.2.2 Virtual Objects (V_{OBJ})

For administering the JA task, we needed virtual objects (V_{OBJ}) that were presented on the shelves inside the virtual room (see Fig. 7). Also, we wanted to use neutral objects to remove any object-related confound in the JA task. These objects were chosen from The Bank of Standardized Stimuli (BOSS) dataset [12]. Google SketchUp [http://www.sketchup.com] was used for designing V_{OBJ} and exported to the Vizard platform.

2.2.3 Design of Intelligent Prompting Protocol Engine (IPPE)

For administering the JA tasks, the agent stood at the center of the virtual room and issued prompting cues towards V_{OBJ} for the user. The audio-visual prompting cues differed in the amount of information offered to the user. Each JA task had four options for JA prompting cues (P1–P4) (see Table 1). The P1–P3 were solely agent-mediated and P4 had a combination of agent and V_{OBJ}. Each of the four agents (see Sect. 2.1.3) administered one of the four JA tasks (JA Task1 to JA Task4).

To facilitate execution of JA tasks, our computer-based system utilized an intelligent prompting protocol engine (IPPE). The IPPE was a state-machine (see Fig. 8)

Table 1 Types of JA prompting cues

JA prompting cue ID	Information content
P1	"look at that" + gaze cue (*look left/look right*)
P2	P1 + head turn (*left turn/right turn*)
P3	P2 + finger pointing (*pointing left/pointing right*)
P4	P3 + V_{OBJ} mediated sparkling

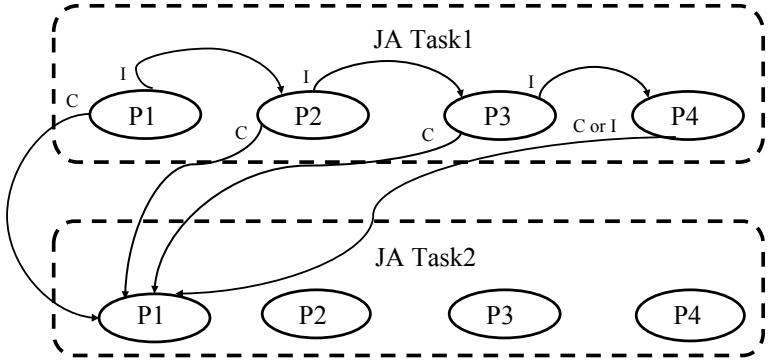


Fig. 8 Intelligent task switching from JA Task1 to JA Task2 (*Note* I = Incorrect response; C = Correct response to JA prompting cue; P1, P2, P3, P4 = JA prompting cues)

based engine that autonomously decided on the type of JA cue (P1–P4) to be offered to a participant based on one's ability to decipher a cue. If a participant was in JA Task1 and was able to pick up P1, then, our system offered JA Task2 with another agent prompting P1. Else, for an incorrect response (or no response for a pre-defined duration (25 s taken as an initial approximation), the system remained in the same JA task while the agent offered the next prompting cue with increased information. For example, if a participant was in JA Task1 and was not able to pick up P1, then the agent offered P2 and this continued till P4 that was in turn followed by the JA Task2. This continued till all the JA tasks were completed.

3 Materials and Methods

Once the agents and agent-mediated task were designed, our next goal was to understand whether the user perceived the agent's face as was thought by the designer and their application in computer-based agent-mediated JA study. For this, we carried out a survey (for faces) and a computer-based agent-mediated JA study.

3.1 Participants

Volunteering participants were recruited both for the survey questionnaire used for validation of the agents' faces and for the computer-based agent-mediated JA study.

Table 2 Participant's characteristics

ID	Age (yrs.) mean (SD)	SRS[a] mean (SD)	SCQ[b] mean (SD)	C-JARS[c] mean (SD)
Group_TD (N = 6)	6.17 (0.41)	47.17 (1.94)	4.17(1.16)	3.37(0.21)
Group_ASD (N = 6)	6.33 (0.52)	72.5(5.68)	18.8(1.722)	−1.2(0.494)

Note N = number of participants, [a]cut-off = 59; [b]cut-off = 15; [c]range = (−4 to +4), SD: standard deviation, yrs. = in years

3.1.1 Volunteers for Validation of Agents' Faces

For this, we recruited 41 typically developing (TD) volunteers [Mean (SD) = 7.6 (1.6) years] from neighboring schools. All of them were comfortable in all or some of the languages (Hindi, English, and Gujarati) used for the survey. Subsequently, once these agents were validated by the volunteers, we used these agents for a JA task-based application. Here, we designed a small study (computer-based agent-mediated JA study *henceforth*) in which both TD children and those with ASD participated (closely age-matched with our volunteers of the survey).

3.1.2 Participants of Computer-Based Agent-Mediated JA Study

A total of 12 participants with 6 participants with high-functioning ASD (Group_ASD *henceforth*) and control group of 6 TD individuals (Group_TD *henceforth*) volunteered. Both groups were age-matched. While enrolling the participants, we considered the Social Responsiveness Scale (SRS) [13] and Social Communication Questionnaire (SCQ) [14] for autism measures. Also, we used the Childhood Joint Attention Rating Scale (C-JARS) [15] which gives the quantification of JA skill in an individual in daily living. The average SRS and SCQ scores indicated that those with ASD were above the clinical thresholds (cut-off score) that were not the case with Group_TD participants (see Table 2). Again, the C-JARS scores were mostly < 0 for Group_ASD (in contrast to the Group_TD for whom the scores were > 0) indicating that the Group_ASD had JA-related milestones.

3.2 Procedure Followed During Computer-Based Agent-Mediated JA Study

The agent-mediated JA task session required a commitment of ~ 20 min from each participant. Once a participant entered the study room, the experimenter introduced herself to the participant and asked him/her to sit on a chair. Then the experimenter gave a brief overview of the task using a visual schedule and told that they would

meet virtual characters (that they were asked to think as virtual peers) during the JA tasks. This was followed by the signing of the consent form by the caregivers. All the participants were told that they were free to leave the study at any point if they felt uncomfortable. The task began with the agent appearing in a virtual room (Sect. 2.2.1) with the head oriented downwards (*downside turn* (see Sect. 2.1.4)) followed by head and gaze orientation being *no turn* and *look straight* to establish eye contact. Subsequently, the agent delivered audio-visual prompting cues (see Table 1) to administer the JA tasks to which the participants responded with a mouse click. After the participants finished interacting with our computer-based agent-mediated JA study, the experimenter carried out a post-study feedback survey.

3.3 Post-study Feedback Survey

The purpose of carrying out this survey was to ensure whether our agent-mediated environment used for the JA task was acceptable to the participants. For this, we framed 5 questions (Q1–Q5) rated on a 5-point scale (with 0 being "not at all" and 4 indicating "very much"). The questions Q1 (Did you like the virtual peer?) and Q2 (Did you feel that the virtual peer looked like real human being?) were aimed to understand the participant's impression on the agent administering the JA tasks. The question Q3 (Did you enjoy interacting with your virtual peer?) and Q4 (Are you interested in interacting with these virtual peers in future?) was meant for understanding their views on their interaction with the agent during the JA task. Finally, the last question (Q5 (Did you feel comfortable with the JA tasks?)) was asked to understand whether the agent-mediated JA task was acceptable to the participants.

4 Results

Having designed the agents, we wanted to understand how the user perceived the agents' faces. For this, we conducted a survey (see Sect. 2.1.3.1) and here we will present our observations based on the users' responses. Our research involved applying these agents in computer-based JA tasks. Thus, we wanted to understand the acceptability of our agent-mediated JA tasks by our participants and here we will present the results of our post-study feedback survey (see Sect. 3.3). Finally, given the JA tasks, we wanted to study whether our system can quantify the capability of our participants to pick up JA prompting cues (by gradually increasing information content based on individualized capability and delivered by an agent).

4.1 Understanding the User's Perception of the Agent's Face

We wanted to select faces (for our agent-mediated JA tasks) that had attributes of neutral expression and Indian look. For this, we conducted a survey (Sects. 2.1.3.1 and 3.1.1) on designed faces of the agents. Based on the users' responses to the survey questions (Q1 and Q2), we computed the $\%\text{Recognition}_{\text{FACE}}$ using (1) for two attributes namely, agent's neutrality of expression and nationality.

$$\%\text{Recognition}_{\text{FACE}} = \frac{\mathbf{NCR_{FACE}} - \mathbf{NIR_{FACE}}}{\mathbf{total\ no.of respondents}} \tag{1}$$

where NCR_{FACE} = number of times the attributes of a face was correctly recognized, NIR_{FACE} = number of times the attributes of a face were incorrectly recognized.

Table 3 presents our findings on the agents' faces (M1–M4, F1–F4) while considering two aspects i.e., neutrality and nationality (Indian) expressed in terms of $\%\text{Recognition}_{\text{FACE}}$ based on the users' responses. It can be seen from Table 3 that 6 out of 8 faces were rated as having a neutral expression $\geq$ 95% of the time and 5 out of 8 were rated as having an Indian look 100% of the times. We chose M1, M3, F1 and F4 faces (i.e., selected faces; see Table 3) that scored a minimum of 95% on the neutrality aspect and 100% on the Indian look. Please note that such thresholds were taken as an initial approximation for our study.

Table 3 Survey results for the agents' faces

Agent ID	Face	Neutrality (%)	Nationality (% Indian)	Agent ID	Face	Neutrality (%)	Nationality (% Indian)
Ml		100	100	F1		100	100
M2		95	95	F2		90	95
M3		95	100	F3		90	100
M4		100	95	F4		95	100

4.2 *Understanding the Acceptability of Our Agent-Mediated JA Tasks*

After the participants completed the computer-based agent-mediated JA study, we collected post-study feedback survey to understand the acceptability in terms of their impression with regard to (i) the agents, (ii) their interaction with the agents and (iii) the agent-mediated JA tasks. From Fig. 9, we can see that most of the TD children and children with ASD had a positive impression about the selected agents.

As far as response to Q1 was concerned, 5 out of 6 participants in each of Group_TD and Group_ASD rated all the selected faces as 4 (scale of 0–4 with 0 and 4 indicating "not at all" and "very much" respectively) except that for M2 and F2 (that were rated as 3). The reason was that they liked the faces M1 and F1 more than M2 and F2. As far as response to Q2 was concerned, mostly the participants could relate the agents with their acquaintances. Only 2 participants in each of Group_TD and Group_ASD felt that the hair on the agents looked more black than usual though none expressed that these did not look real. From the response to Q3, we could see that both the groups enjoyed their interaction with the agents. In fact, most of them expressed that they felt like interacting with their peers and brothers/sisters. As far as future interaction possibilities (Q4) was concerned, all the participants of Group_ASD expressed their willingness to interact with our agents in the future. Again, the ratings by all the TD participants (except one who awarded 3 points) were similar to that of the Group_ASD. The single TD participant who gave 3 points expressed that he was not sure whether he would like to get another exposure in the future. When asked about their impression on their interaction with the agent-mediated JA tasks (Q5), most of the participants mentioned that they were quite comfortable in interacting with our JA tasks. However, 3 and 2 participants from Group_TD and Group_ASD respectively, said that they felt somewhat less comfortable while using mouse clicks for responding to the JA cues.

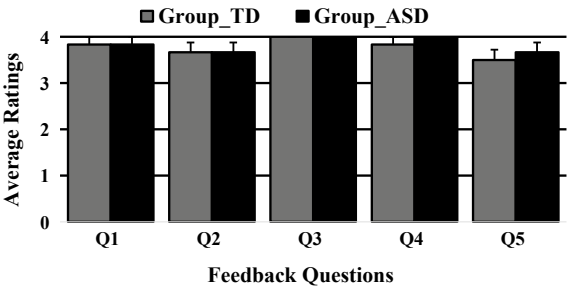


Fig. 9 Average ratings for post study feedback questions by both groups

4.3 Comparative Group Analysis of the Capability to Pick up Agent-Mediated Prompting Cues

Given the fact that (i) children with ASD have difficulties in picking up JA cues [3] and (ii) the evidence from literature indicating the potential of using JA cues (with gradual increase in the prompting information) as practiced by the interventionists in conventional settings [9], we wanted to understand the applicability of our agent-mediated JA tasks for administering JA skill training. Also, we wanted to understand the comparative estimates of one's capability of picking up JA prompting cues (as offered by our intelligent prompt protocol engine) for both Group_TD and Group_ASD.

For this, we calculated the distribution in terms of % of the number of participants in Group_TD and Group_ASD who have been successful in picking up each prompting cue (P1–P4; see Sect. 2.2.3). For Group_TD, all the participants were able to understand and decipher gaze-based cue (P1) and thus were not offered any other JA prompting cue (See Fig. 10). From this, we might infer that it was quite easy for TD participants to decipher the gaze cue (P1) delivered by our agents. For the Group_ASD, observations were different from Group_TD. In fact, none of the participants from Group_ASD were able to decipher P1. Such an observation can be possibly attributed to the fact that individuals with ASD tend to avoid the face region of a communicator [16]. This caused our agent-mediated system to increase the amount of cueing information to P2. However, only 2 participants out of 6 were able to pick up P2. Added with finger pointing, the P3 could elicit increased correct responses from the Group_ASD leaving behind a smaller group of participants with ASD to be offered with the P4 (See Fig. 10).

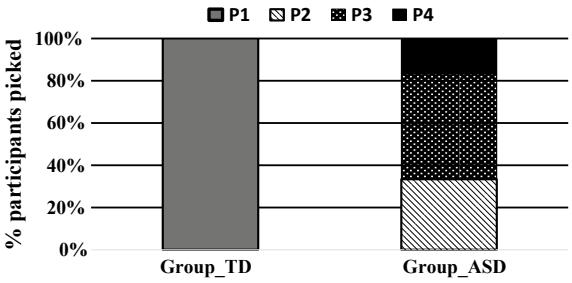


Fig. 10 Groups analysis of JA prompting cues picked by both groups (*Note* P1, P2, P3, P4 are JA prompting cues)

5 Conclusion and Future Work

In our present study, we have offered the steps required to design virtual characters as artificial agents (agents) that looked like Indians. Specifically, our agents were designed to have complexion and texture of hair, skin, and color of eyeballs, etc. so as to have an Indian look. We selected four agents (2 males and 2 females) from a database of 8 designed agents based on a survey conducted among volunteers. Thus, the first contribution of our presented research is the design of a database of agents that can offer 3-D humanoid embodiment, looks Indian and capable of demonstrating animated moves that are naturalistic. These selected agents were used as facilitators that were capable of delivering varying prompting cues for administering computer-based JA tasks. The second contribution of our presented research was augmenting the agents with an intelligent prompt protocol engine (IPPE) thereby making it autonomous and adaptive to individualized JA capabilities of each child with ASD. Specifically, the IPPE was aimed to make the agents capable of understanding one's ability to understand and decipher a prompting cue. If a participant was not successful in picking up a prompting cue, then the IPPE triggered the agent to deliver a different prompting cue that was more informative than that delivered earlier. Such adaptive delivery of prompting cues, an essential part of JA skill training was embedded in our designed computer-based agent-mediated JA task platform that was the third contribution of our present work. Age-matched TD children and children with ASD participated in our computer-based agent-mediated JA study. Results indicate that our agent-mediated JA tasks were acceptable by our participants. Also, the results indicate that our agent-mediated JA tasks can quantify the capability of our participants to pick up JA prompting cues.

Although the results of our study are positive, yet, with certain limitations still remain. For example, we had a limited number (four JA tasks with each task having four prompting cues (P1–P4)) of JA tasks. In the future, we plan to design more tasks for carrying out a full-fledged intervention study. Another limitation of our present work is the sole use of the 3-D humanoid agent for administering the computer-based JA tasks. Given interesting findings reported in the literature on the use of non-human agents as mediators of JA skill training for children with ASD [2], it might be critical to carry out our study with both humanoid and non-human like agents as mediators of JA skill administration. In the future, we plan to extend our present study by incorporating non-human agents followed by a comparative analysis of the implications of humanoid versus non-human agents on the JA skill training of children with ASD. Still, another limitation was the medium available to the participants to respond to the prompting cues. Specifically, the participants could use only a computer mouse for giving responses that might be an issue for participants with low-functioning ASD. In the future, we plan to use a touch-sensitive computer monitor to alleviate this problem.

Additionally, though our present study shows the acceptability and the feasibility of using computer-based agent-mediated JA task administration for children with ASD, yet, there is a lot of scope for extending the present study. Specifically, in the

future, we plan to extend our agent-mediated JA task to develop JA skill intervention modules. Since, in our present research, the aim was to design a computer-based agent-mediated JA task platform, we wanted to carry out a proof of concept study involving a limited number of participants so that we can understand the acceptability and the feasibility of application of such a platform for children with ASD. Thus, in the future, we plan to continue this study with a bigger participant (with ASD) pool. Since we run collaborative endeavors with neighboring special needs schools that offer intervention services to children with ASD, carrying out such extended studies is feasible. These, in turn, can be deployed as supportive tools in the hands of the clinicians thereby contributing to JA skill training among individuals with ASD.

Acknowledgements The authors are thankful to the participants, from local regular school and BM institute of mental health for helping us to enroll the participants for our study. Also, we thank the TCS research fellowship and IIT Gandhinagar for supporting this research.

References

1. Gaggioli, A., Mantovani, F., Castelnuovo, G., Wiederhold, B., Riva, G.: Avatars in clinical psychology: a framework for the clinical use of virtual humans. Cyberpsychology Behav. **6**(2), 117–125 (2003)
2. Bernardini, S., Porayska-Pomsta, K., Smith, T.J.: ECHOES: An intelligent serious game for fostering social communication in children with autism. Inf. Sci. **264**, 41–60 (2014)
3. Mundy, P., Jarrold, W.: Infant joint attention, neural networks and social cognition. Neural Netw. **23**(8–9), 985–997 (2010)
4. Lu, A., Chan, S., Cai, Y., Huang, L., Nay, Z.T., Goei, S.L.: Learning through VR gaming with virtual pink dolphins for children with ASD. Interact. Learn. Environ. **26**(6), 718–729 (2018)
5. Caruana, N., Stieglitz Ham, H., Brock, J., Woolgar, A., Kloth, N., Palermo, R., McArthur, G.: Joint attention difficulties in autistic adults: an interactive eye-tracking study. Autism **2**(4), 502–512 (2018)
6. Dalwai, S., Ahmed, S., Udani, V., Mundkur, N., Kamath, S.S., Nair, M.K.: Consensus statement of the Indian academy of pediatrics on evaluation and management of autism spectrum disorder. Indian Pediatr. **54**(5), 385–393 (2017)
7. Wing, L.: Language, social, and cognitive impairments in autism and severe mental retardation. J. Autism Dev. Disord. **11**(1), 31–44 (1981)
8. Plaisted, K.C: Reduced Generalization in Autism: An Alternative to Weak Central Coherence (2015)
9. Taylor, B.A., Hoch, H.: Teaching children with autism to respond to and initiate bids for joint attention. J. Appl. Behav. Anal. **41**(3), 377–391 (2008)
10. Itti, L., Dhavale, N., Pighin, F.: Realistic avatar eye and head animation using a neurobiological model of visual attention. In: Applications and Science of Neural Networks, Fuzzy Systems, and Evolutionary Computation VI, vol. 5200, pp. 64–79. International Society for Optics and Photonics (2003)
11. Babu, P.R., Lahiri, U.: Understanding the role of Proximity and Eye gaze in human–computer interaction for individuals with autism. J. Ambient. Intell. Hum. Ized Comput. 1–5 (2019)
12. Brodeur, M.B., Dionne-Dostie, E., Montreuil, T., Lepage, M.: The Bank of Standardized Stimuli (BOSS), a new set of 480 normative photos of objects to be used as visual stimuli in cognitive research. PLoS ONE **5**(5), e10773 (2010)

13. Coon, H., Villalobos, M.E., Robison, R.J., Camp, N.J., Cannon, D.S., Allen-Brady, K., Miller, J.S., McMahon, W.M.: Genome-wide linkage using the social responsiveness scale in Utah autism pedigrees. Mol. Autism **1**(1), 8 (2010)
14. Chandler, S., Charman, T., Baird, G., Simonoff, E., Loucas, T.O., Meldrum, D., Scott, M., Pickles, A.: Validation of the social communication questionnaire in a population cohort of children with autism spectrum disorders. J. Am. Acad. Child Adolesc. Psychiatr. **46**(10), 1324–1332 (2007)
15. Mundy, P., Novotny, S., Swain-Lerro, L., McIntyre, N., Zajic, M., Oswald, T.: Joint-attention and the social phenotype of school-aged children with ASD. J. Autism Dev. Disord. **47**(5), 1423–1435 (2017)
16. Trepagnier, C., Sebrechts, M.M., Peterson, R.: Atypical face gaze in autism. Cyber Psychol. Behav. **5**(3), 213–217 (2002)

Automated Assessment of ER Model Using the Domain Knowledge

Muhammad Javed and Yuqing Lin

Abstract It is a challenging task to develop a complete and correct Entity Relationship Model (ERM) from the requirements. Quality of artifacts in the later stages of software development (e.g. logical database design, physical database design, and the final product) dependents on the conceptual models. Domain Knowledge (DK) plays a key role while extracting the artifacts from requirements. It is agreed upon that most errors in the early stages of software development are due to the lack of adequate DK. In this paper, we are proposing an automated assessment approach, which focuses on some major issues of ERM such as completeness and correctness. The automatically extracted ERM is used as a reference for the assessment of a manually developed model from the same set of requirements. We trained the Skip-gram model of word2vec for extracting the DK, which is used for assisting in errors detection and ERM's labels matching. As a case study, we considered models from the business domain. Inputs of this automated technique are reference model, DK, and the model to be evaluated. The output is a list of errors (indicating the sources) and suggestions. The results show the proposed approach is having a noticeable improvement over the existing approaches.

Keywords Conceptual diagram · Entity relationship model · Automatic quality assessment · Domain knowledge

1 Introduction

Extracting ERM from the requirements is an important step in the Software Development Life Cycle (SDLC) [1–3]. Quality of database design and software depends on the quality of ERM [4, 5]. Traditionally, the quality of conceptual models is assessed at later stages of software development [6]. However, the cost of fixing errors

M. Javed (✉) · Y. Lin
School of Electrical Engineering and Computing, The University of Newcastle, Newcastle 2308, Australia
e-mail: m.javed@uon.edu.au

Y. Lin
e-mail: yuqing.lin@newcastle.edu.au

R. Lee (ed.), *Computer and Information Science*, Studies in Computational Intelligence 849, https://doi.org/10.1007/978-3-030-25213-7_10

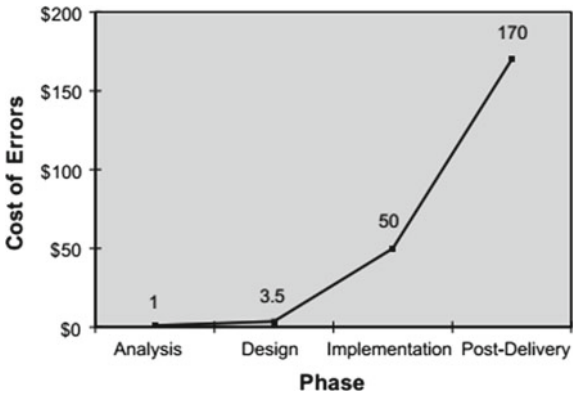


Fig. 1 The relative cost of removing errors [8]

increases as the project matures [7]. Barry Boehm is one of the pioneer researchers who calculated the cost of defect removal during the phases of SDLC [8]. According to Boehm, finding the same error in the design phase will cost approximately 3.5 times more as compared to the analysis phase. Many researchers and practitioners mentioned that the expense of error detection during the later stages is very high compared to the detection at early stages of SDLC [9–12] (Fig. 1).

Assessment of the ERM is an important task in software development [13]. A number of techniques have been proposed by the researchers to evaluate the ERMs [13–21], but to the best of our knowledge, little effort has been made to integrate the DK in evaluation.

In this paper, we are presenting an automated assessment process to discover the errors in relation to the completeness and correctness of ERMs. There are several existing tools for ERM extraction. The recent one is presented by Javed and Lin in [22]. ERM extracted using the automated technique is used as an reference for the verification of errors found (by applying the DK) during the assessment of the manually derived model from the same set of requirements. The reference model is also used to evaluate some of the completeness related issues of manual extracted ERM. We are not suggesting that the auto-extracted ERM is better than the manually derived one, rather we are suggesting that all the relationships and information that an auto-extraction process has learned from the requirement document should be included in some ways in the manually extracted ERM.

The assessment process contains two components: Structural Analysis (StrA) and Semantic Analysis (SA). StrA is to evaluate the models in order to find the correctness and completeness related errors. SA is the application of DK. The assessment requires to understand the similarity between the labels of ERM (i.e. names of entities, attributes, and relationships). For example, "*client*" and "*customer*" are domain-based synonyms which may be used interchangeably due to the inconsistency in the requirements document. SA is an umbrella activity (continue throughout the StrA). DK applies not only to match labels but also for the detection of errors in ERM.

In the remainder of this paper, Sect. 2 is the detailed literature review. Section 3 contains the list of errors and NLP based challenges we focus on. A detailed explanation of the proposed approach is in Sect. 4 while Sect. 5 contains a case study. Section 6 is to present the preliminary evaluation of the accuracy and efficiency of the proposed approach. Conclusion and future works are presented in Sect. 7.

2 Literature Review

Assessment depends on quality metrics. The researcher uses the following metrics as checklists/measurement goals while evaluating the quality attributes of conceptual models.

In [14], Moody and Shanks proposed a framework to evaluate the quality of both product and process. Stakeholders, quality factors, weighting, quality metric, improvement strategy, and review checkpoint are the building blocks of the framework. The role of stakeholders is to find the root causes of quality problems. Authors considered eight quality factors i.e. completeness, integrity, flexibility, understandability, correctness, simplicity, integration, and implementation for the evaluation of data models. Weighting is for defining the importance of quality factors. Moody defined 29 metrics as a checklist while evaluating the quality of data models. Each quality factor of data models is evaluated by considering its corresponding metric. The authors present the initial version of quality metrics in [15]. Moody presented a survey [4] on the quality of conceptual models. Hussain [21] tried to revise these quantitative metrics. He considered the metrics explained by several researchers i.e. proposed by Moody and Hussain et al. He has defined quality metrics to evaluate the completeness, correctness, and complexity of ERM. Krogstie [23] considered quality metrics to assess the quality of data models based on organizational semantics. He tried to measure the empirical quality (connections between nodes), syntactic quality (number violation to data modeling standards), pragmatic quality (understandability), Social quality (agreements) and deontic quality of the conceptual model. Kesh [24] presented a framework for the validation and verification of ERM. He considered soundness, consistency, and conciseness w.r.t the structure while completeness, cohesiveness, and validity w.r.t the content, to evaluate the quality of ERM. Completeness of the model is calculated by combining the score of its components. Overall quality score is calculated from the weights of usability, maintainability, accuracy, performance and the scores of behavioral factors. Gosain et al. [25] proposed metrics to compute the structural complexity (maintainability) of conceptual models for the data warehouse. Authors proposed metrics to evaluate the dimensional hierarchy of the model. Metrics used to count the hierarchies, alternate paths in hierarchies, level, and a number of dimensions of hierarchies.

Thi and Helfert [26] proposed a framework for the evaluation of an information system's quality by considering the user perspective (i.e. evaluation of final product) and developer perspective (i.e. modules implemented). Authors discussed the quality

factors and their interdependency for the phases of the software development life cycle.

Thomas et al. [27] presented a technique for the automated assessment of students' ERM for marking. In this effort authors focused on labeling error (misspelling) by calculating the edit distance and weights. Edit distance of two strings is calculated by comparing and also weights (0–1) are assigned to the components (depending on importance) of ERD. The similarity of the two diagrams is the weighted sum of components. Marks assigned to the student diagrams based on its weighted similarity values with which it compared (i.e. reference diagram). If the matching result of labels is greater than the threshold i.e. 0.5 then one mark, otherwise no mark awarded.

Batmaz and Hinde [28] presented a semi-automated technique for the assessment of database diagrams. In this technique, authors assessed the group of components instead of a complete diagram. Entities having the same name and the same number of attributes considered matched. Authors used Latent Semantic Analysis (LSA) for the matching of text.

Hussain et al. [29] presented an approach to measure the completeness of ERMs by calculating the Fuzzy Completeness Index. ERD is transformed into the graph and functional dependencies calculated by defined rules. Finally, a fuzzy logic based approach applied to the graph to measure the completeness of ERD.

Lino and Rocha [30] tried to evaluate students' ERD. It is a supervised machine learning based approach. Training data is generated by converting ER diagrams into SQL metrics. Different learning algorithms are tested and the best one is selected by comparing mean square error and determination coefficient.

Choppella et al. [31] presented a technique for the validation of ERM's correctness using the specification language PVS (based on set theory) and theorem proving. Specification language is used to model a sample database with different level of abstraction. Tables of the model are implemented by defining a specific structure. Machine checkable logical notations used to validate the correctness. Data models improved by continuous user reasoning and interactively property proving about the specification.

Kazi et al. [32] proposed an automated technique for the semantic evaluation of ERM by comparing domain ontology with the elements of data models.

In these efforts, mostly authors tried to validate the ERM by focusing on the quality metrics defined but without much consideration of the DK. According to Schenk et al. in [33] good molders mainly rely on their knowledge gained through the experience. Conceptual modeling and its evaluation require the DK. Especially in data modeling, it helps to find the redundancy, anomaly, and inconsistency based errors w.r.t the domain. More efforts needed for the domain based semantic analysis of ERMs.

3 Common Errors in ER Modeling

3.1 Structural Errors of ERM

StrA is to find errors by comparing the components (entities, attributes, relationships, and cardinalities) of the manual model to those in the reference model. Our main focus is to evaluate two of the most important categories of error i.e. completeness and correctness.

3.1.1 Correctness

Is to know "*Does the model confirm rules of the data modeling technique?*" This includes complying with diagramming conventions, naming rules, definition rules and rules of composition and normalization [14, 15, 23, 34, 35]. Following errors types are related to the correctness of ERM.

(a) ***Redundancy***: Data redundancy might occur in ERMs due to the repetition of tuples. It is a common and a serious issue in the database systems. Very often the redundancy of data is the source of inconsistency in the software. Similarly, data redundancy leads to data anomalies and corruption. Therefore, it should be avoided when creating an ERM [14, 34]. The following redundancy based errors are addressed in this approach.

- Entities and attributes with names that are domain-based synonyms e.g. "*Customer*" & "*Client*".
- Entity within entity (intra-entity redundancy [34]) e.g. *Sale (Sale_ID, Sale_Date, Name, Address, City)*
- Repeating attributes within the entity (intra-entity redundancy [34])
- Repeating attributes in different entities (inter-entity redundancy [34]) e.g.

  Department (ID, department_name)
  Instructor (ID, name, department_name, salary)

(b) ***Suitability***: The model should have a correct structure [24]. Here are the typical errors:

  - Hypernym/abstract entities. e.g. "*Employee*" instead of "*Administrator & Salesman*"
  - Attributes of an entity scattered in other entities.
  - Attribute been mistakenly modeled as an entity or vice versa e.g. "*order_status*" is an entity that is modeled as an attribute.
  - The entity is mistakenly modeled as a relationship e.g. "*gerund*"
  - Wrong selection of maximum and minimum cardinality.
  - Irrelevant attribute e.g. "*DOB*" in "*invoice*" entity.

3.1.2 Completeness

Refers to, whether the model contains all information to support the required functionality of the system. The following errors will lead to an incomplete model [14, 15, 23, 26, 34, 35].

- Missing components i.e. Entities, Attributes and Relationships
- The entity having only one attribute.
- Extra artifacts. Items included in the model that do not correspond to any user requirement e.g. "*System*".

3.2 Natural Language Processing Based Challenges

Along with the above-mentioned structural errors, we also tried to address some NLP based challenges when automatically assessing manually prepared diagrams. A manually extracted diagram might contain any of the following difficulties.

- *Symbols*: special characters in the names of entities, attributes & relationships e.g. *Employee_ID* & *Employee ID* [36].
- *Misspelled*: labels in the diagrams could have spelling mistakes or cultural language differences e.g. "*colour* & *color"*, "*labour* & *labor*" [27, 36, 37].
- *Concatenation*: words might be connected without any space e.g. "*Generate Report*" as "*Generatereport*" [36].
- *Domain based Synonyms*: model might contain the domain-based synonyms e.g. "*unit_cost* & *price*", "*Username* & *ID*" [27, 36, 37].
- *Hypernym*: either the reference model or the diagram for validation may contain the generalized forms of entities e.g. both "*Manager* & *Worker*" in one diagram might be represented as "*Employee*" in other [36].
- *Abbreviations*: Instead of a complete phrase, there could be an abbreviation [36].
- *Form of Verb*: a different form of the verbs could appear in both diagrams e.g. "*Manage* & *Managed*", "*Delete* & *Deleting*".
- *Singular* & *Plurals*: This problem may arise e.g. "*Apply* & *Applies*", "*Library* & *Libraries*".
- *Length of Phrases*: words or phrases representing a similar artifact of the diagram could have a different number of characters e.g. "*Create New Account*" & "*Create Account/Signup*" [36, 37].

4 Proposed Approach

In this section, we discuss the proposed approach in detail. The inputs to the process are the ERMs (mapped to text format) and DK while the output is a list of errors and suggestions. For the validation of proposed approach, a prototype has been developed. Following is the flow chart of the proposed technique (Fig. 2).

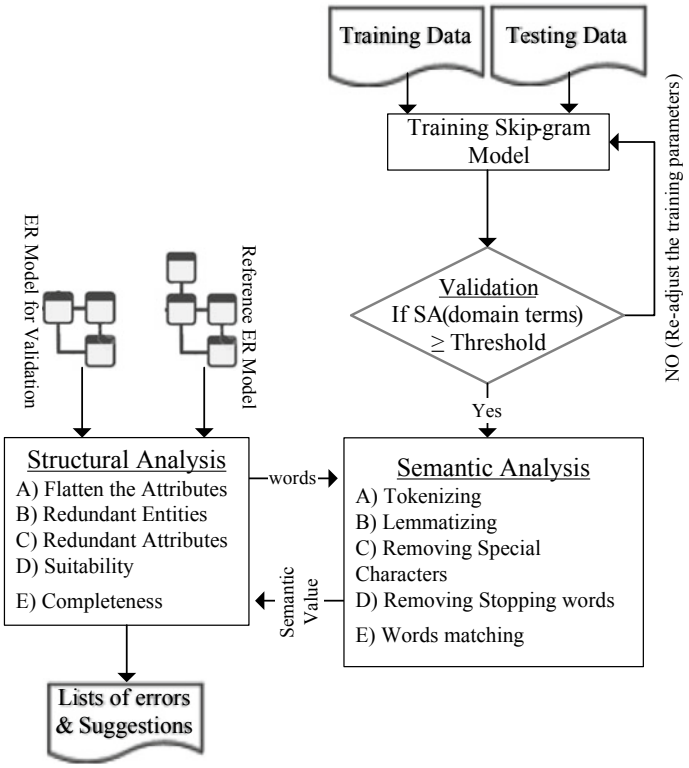


Fig. 2 Flow chart of the proposed approach

4.1 Domain Knowledge

The first objective for utilizing DK is to find structural errors by considering the semantic similarity values of the components. Semantic similarity value represents how often two components appear together in the documents. If semantic similarity value of two components is higher than a threshold it means there is some sort of relation between them e.g. entity-attributes or relationship (if both are entities), etc. For instance, if the semantic similarity value of an attribute "*price*" is higher with the entity "*Product*" and lower with "*Employee*" so, it could be the attribute of "*Product*" (hypothesis considered during the irrelevant attributes detection process).

Secondly, SA is for the matching of labels (i.e. Names of entities, attributes, and relationship) using the DK. A conceptual diagram might contain the domain-based synonyms due to the inconsistency in the requirements document. Therefore, an efficient SA technique is required to match the labels while performing the StrA. Researchers used Latent Semantic Analysis (LSA) and WordNet based labels matching techniques [38]. One of the efficient methods of finding contextual similarity of words is the use of word embedding methods. Performance of the LSA is better when the training data is small but for the large dataset, word2vec performs effi-

Table 1 Sample results of business domain based semantic analysis

Business terms examples	Semantic analysis
Client & Customer	0.97
Register & Enroll	0.92
Cash & Card	0.95
File & Record	0.94
Item & Product	0.95

ciently [39]. The most effective concept of word2vec is the concurrence of the word vectors, not just to find the meaning but also for the same context. In [39] authors conducted a comparative study on word embedding methods (LSA, Word2Vec and Glove) and proved that word2vec (skip-grams with negative sampling) is the more efficient model for the word vector representation from the text data. After analyzing the performance of existing techniques, we decided to train the Skim-gram model of word2vec introduced by [40]. We also used Stanford CoreNLP 3.8 API for the preprocessing of words.

4.2 Training of Neural Network

Selection of the training data is equally important as the selection of training algorithm. It is a challenging task, depending on the required functionality to be performed [38]. We have selected business applications as our case study, therefore, we collected the business and financial news data dumps of Reuter News and Bloomberg News for the training of Skip-gram model. We used the Stanford CoreNLP 3.8 API for the preprocessing of data. After removing the stopping words, we identified 37 billion tokens. Training and testing data is separated based on a silver bullet rule of 80:20 (80% training data and 20% testing data). We trained the skip-gram model by using the various combinations of parameters i.e. window size 5–10, dimensions 100–1000 and epochs 3–15. After each training process, testing data is used to validate the efficiency. This process continued until we got the semantic results (of most frequent domain terms) greater than the threshold value. To get the most frequent terms, we calculated the frequency of terms and selected the top 30%. Finally, we decided with window size = 5, Dimensions = 600, and epochs = 3. Some of the results of business domain-based SA are listed in Table 1. The trained neural network is used for the matching of labels in SA while finding the structural errors.

4.3 Mapping of ERM to the Text Format

To convert the ERMs in textual format, the same template used as in [22]. The textual template for ERM is as below.

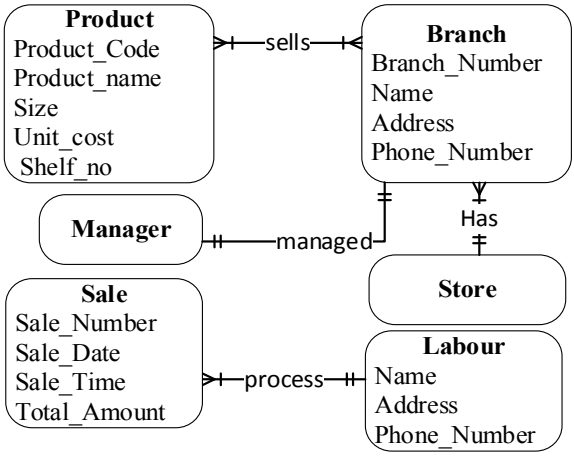


Fig. 3 Sample ERM for mapping

Entities with attributes

***Entity*$_1$** *(Attributes1, Attributes*$_2$*,… Attributes*$_n$*)*

***Entity*$_2$** *(Attributes1, Attributes*$_2$*,… Attributes*$_n$*)*

Relationships

Entity$_1$ *> Cardinality (Relationship) Entity*$_2$ *> Cardinality*

For Example, following is the textual representation of ERM depicted in Fig. 3.

Mapping of Entities and Attributes in Text format

Manager

Store

Product*(Product_code, Product_name, size, unit_cost, shelf_no)*

Branch *(Branch_Number, name, address, Phone_number)*

Sale*(sale_number, date, time, total_amount)*

Labour *(name, address, telephone)*

Mapping of Relationships in Text format

store > 1,1 (has) branch > 1,n

branch > 1,1 (managed) manager > 1,1

branch > 1,1 (sells) product > 1,n

product > 1,n (sold) branch > 1,1

labour > 1,1 (process) sale > 1,n

4.4 ERM Assessment

Following algorithms are for the assessment of ERMs. First Algorithm is to find the redundancy-based errors while the second is to generate suggestion by considering the errors related to suitability and completeness.

Algorithm I: To detect the redundant artifacts from the ERM to be validated.

Initializations: ERM_V:=Entity Relationship Model for validation, ERM_R:= Reference Entity Relationship Model, $List\_A_V$:= Flattened Attributes from the manual Model (to be validated), $List\_A_R$:= Flattened Attributes from the Reference model,

Output: List of redundant artifacts.

1. Load ERM_V ERM_R from the text files
// generating lists of artifacts and flatten the attributes in the format of "Entity_Name.Attribute_Name"
2. Generate $List\_A_V$, $List\_A_R$ from ERM_V & ERM_R
3. For I:=1 to $List\_A_V$.Count
4. For J:=1 to $List\_A_V$.Count
// finding redundant Entities within the manual ERM
5. If SA(I≠J AND $List\_A_I$.entity, $List\_A_J$.entity) ≥ Threshold
6. Errors_List.add($List\_A_J$)
// Repeating attributes within the entity
7. Else If SA($List\_A_I$.entity, $List\_A_J$.entity) ≥ Threshold
8. If SA($List\_A_I$.Attrib, $List\_A_J$.Attrib) ≥ Threshold
9. Errors_List.add($List\_A_J$)
// scattered & repeating attributes in different entities
10. If SA($List\_A_I$.entity, $List\_A_J$.entity) ≤ Th
11. If SA($List\_A_J$.Attrib, $List\_A_J$.Attrib) ≥ Threshold
12. Errors_List.add($List\_A_J$)
13. End of loop started at step # 4
14. End of loop started at step # 3
// validation of errors (Entity within entity, scattered & repeating attributes in different entities) by comparing with reference ERM
15. For I:=1 to $List\_A_R$.Count
16. For J:=1 to $Error\_List_V$.Count
17. If SA($List\_A_I$.entity, $Error\_List_J$.entity) ≥ Threshold
18. If SA($List\_A_J$.Attrib, $Error\_List_J$.Attrib) ≥ Threshold
19. Errors_List.Remove($Error\_List_J$)
20. End of loop started at step # 16
21. End of loop started at step # 15
22. **Return Errors_List**

The above algorithm is to find the redundant artifacts and scattered attributes from the manually extracted model. The initial steps (1 and 2) are to read the models from the text files and generate the lists by flattening the attributes in the format of "*Entity.Attribute*". Steps 5 and 6 are to find the redundant entities while steps 7–9 are to find repeating attributes within an entity. Steps 10–12 are to find any scattered and repeating attributes in different entities by applying the DK. If any redundant artifacts found they will be added in the error list. The error list is verified at steps 16–19 by comparing with the reference model.

SA is the application of DK which is generated by using the word2vec trained on domain data. On each comparison, labels of the artifacts are passed to SA function to find the semantic similarity. SA returns the semantic distance between the strings. The range of semantic value lies between 0 and 1 (0 = not matched and 1 = exact

match). If the semantic value is greater than or equal to the threshold then these are the potential candidate for the identification process (e.g. here for redundant entities).

Algorithm II: for the assessment of ERM suitability w.r.t correctness and completeness

Initializations: ERM_V:=Entity Relationship Model for validation, ERM_R:= Reference Entity Relationship Model, $List\_A_V$:= Attributes from the manual Model, $List\_A_R$:= Attributes from the reference Model, $List\_R_V$:= Relations from the manual Model, $List\_R_R$:= Relation from the reference Model, MR:= Missing Relationships

Output: List of suggestions to enhance the ERM.

1. Load ERM_V and ERM_R from the text files
   // generating lists of artifacts and the list of attributes will be in the format of "Entity.Attribute"
2. Generate $List\_A_V$, $List\_A_R$, $List\_R_V$, $List\_R_R$, from ERM_V & ERM_R
3. For J:=1 to $List\_A_V$.Count
   // irrelevant attribute
4. If SA($List\_A_J$.entity, $List\_A_J$.attrib) ≤ Threshold
5. Suggestion_List.add($List\_A_J$.attrib)
6. End of loop started at step # 3
7. For I:=1 to $List\_A_V$.Count
8. For J:=1 to $List\_A_V$.Count
   // finding the missing relationships
9. bool is_Exist_relation=false
10. If SA(I≠J AND $List\_A_I$.entity, $List\_A_J$.entity) ≥ Threshold
11. For X:=1 to $List\_R_V$.Count
12. If($List\_R_X$.Contain($List\_A_I$.entity & $List\_A_J$.entity)=false)
13. MR.Add($List\_A_I$.entity, $List\_A_J$.entity)
14. End of loop started at step # 11
15. End of loop started at step # 8
16. End of loop started at step # 7
17. For I:=1 to from $List\_A_R$.Count
18. bool is_Exist_entity := false
19. bool is_Exist_attrib := false
20. For J:=1 to from $List\_A_V$.Count
    // Entity as attribute or vice-versa
21. If SA($List\_A_I$.entity, $List\_A_J$.attrib) ≥ Threshold OR
    SA($List\_A_I$.attrib, $List\_A_J$.entity) ≥ Threshold
22. Suggestion_List.add(Reconsider $List\_A_J$)
23. If SA($List\_A_I$.entity, $List\_A_J$.entity) ≥ Threshold
24. Is_Exist_entity :=true
25. If SA($List\_A_I$.attrib, $List\_A_J$.attrib) ≥ Threshold
26. Is_Exist_attribute :=true
27. End of loop started at step # 20
    // missing or irrelevant entity
28. If is_Exist_entity=false
29. Suggestion_List.add($List\_A_I$.entity missing)
    // missing attribute
30. If is_Exist_attrib=false
31. Suggestion_List.add($List\_A_I$.attrib missing)

```
32.    For J:= 1 to List_R_V
       // Entity as a relationship
33.      If  SA(List_A_I.entity , List_R_J) ≥ 0.95
34.        Suggestion _List.add(Consider List_R_J as entity)
35.    End of loop started at step # 18
36. End of loop started at statement # 3
37. For I:= 1 to List_R_R.Count
38.     For J:= 1 to List_R_V.Count
39.       // wrong cardinalities
40.       If SA(List_R_I.cardinalities ≠ List_R_J.cardinalities)
41.          Suggestion.add(List_R_J.Max)
42.     End of conditional statement started at step # 26
43.   End of loop started at step # 25
44. End of loop started at step # 23
    // validation of missing  relationships
45. For x:= 1 to MR.Count
46.   bool is_Missing_relation=false
47.   For I:= 1 to List_R_R.Count
48.      If(List_R_R.Contain(MR_X))
49.         Is_Missing_Relation:=true
50.    End of loop started at step # 49
51.    If (is_Missing_Relation=true)
52.      Suggestion.add(MR_X Relationship missing)
53. End of loop started at step # 47
54. Return Suggestions
```

The algorithm II is to find possible errors related to suitability (using the DK) and completeness from the manually extracted ERM by comparing it with the reference model. As discussed, the reference model encapsulates the relationships between the entities, but the software engineers might develop the ERM embedding these relationships differently. This algorithm returns the list of suggestions if any. Attributes of an entity will be considered irrelevant if the semantic similarity value is less than the threshold value at steps 3–6, because it means that both don't occur together very often, and by the manual analysis of the documents it is noticed that attributes of the entities normally fall together in the same sentence. Steps 7–16 are to find the missing relationships by applying DK. If the semantic similarity of two entities is greater than or equal to the threshold value, it means that both often occur together then there should be a relationship between them. If no relationship found then add it in suggestions after verifying with the reference model at steps 45–53.

Steps 21 and 22 are to find if there are any entities should be considered as attributes or vice versa. In steps 22–32, we tried to find the missing entities and attributes by comparing it with the reference model. Steps 33 and 34 are to find is there any entities which should be considered as relationships. Steps 37–41 are to find errors related to the cardinalities. The same threshold value is used while applying the DK to compare labels in these steps.

5 Case Study

In this section, we are presenting a walk-through of the proposed approach. We have included an reference diagram generated by the automated technique and a manually extracted diagram to be validated.

5.1 Reference ERM

Following is the textual mapping of reference ERM (depicted in Fig. 5). This model is extracted from the requirements document using the automated technique presented in [22].

Following is textual Mapping of Entities & Attributes of ERM depicted in Fig. 4.

Client *(ID, Company_Name, Contact_Name, Contact_Job_Title, password)*
Employee *(Employee_ID, Last_Name, First_Name, Designation, Report_To)*
Address *(Office_Phone, Mobile_Phone, Fax, EMail, Address, City, State, Country, Postal_Code)*
Invoice *(Invoice_No, Invoice_Date, Amount)*
Product*(Product_No, Product_Name, Description, Unit_Price, Unitin_Stock, Reorder_Level)*
Order *(order_No, Order_Date)*
Order_Status *(Order_Status_Code, Order_status, Status_Description)*
Shipment *(Shipment_No, Shipment_Date, Shipment_Description)*

Relationships

Employee > 1,1 (serve) Client > 0,n
Employee > 1,1 (manage) Employee > 0,n
Employee > 1,1 (has) Address > 1,n
Client > 1,1 (has) Address > 1,n
Shipment > 1,n (deliver) Address > 1,1
Client > 1,1 (place) Order > 0,n
Order > 1,n (has) product > 1,n
Order > 1,1 (has) Invoice > 1,n
Order > 1,n (has) Order_status > 1,1
Order > 1,1 (posted) Shipment > 1,1

5.2 ERM for Validation

Following is the textual mapping of entities & attributes of ERM depicted in Fig. 5.

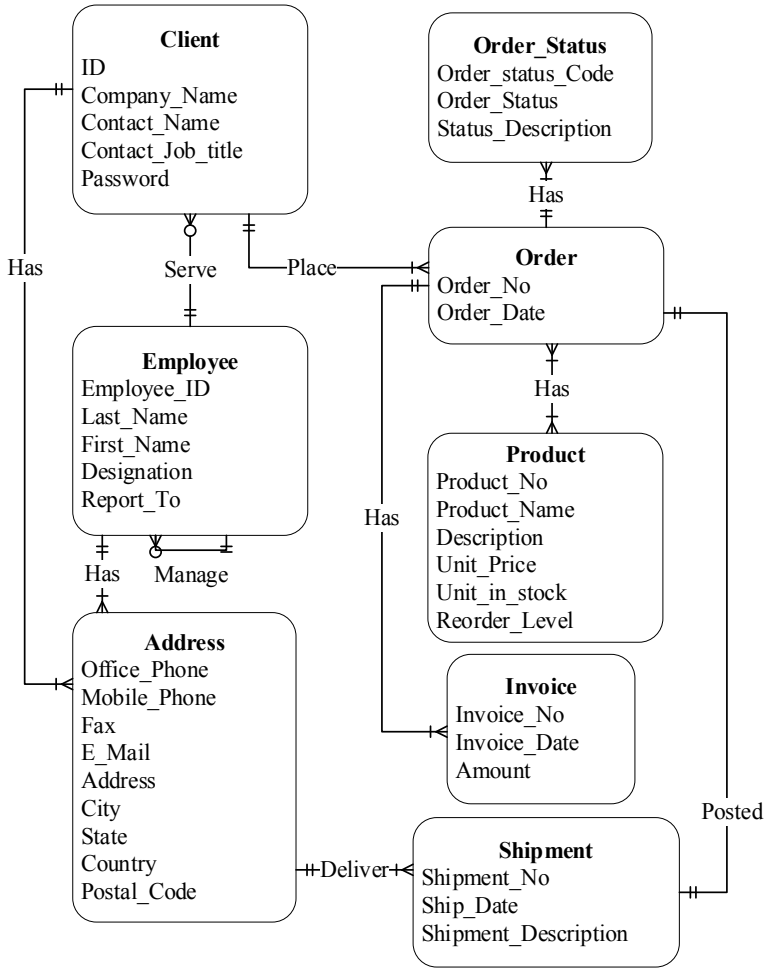
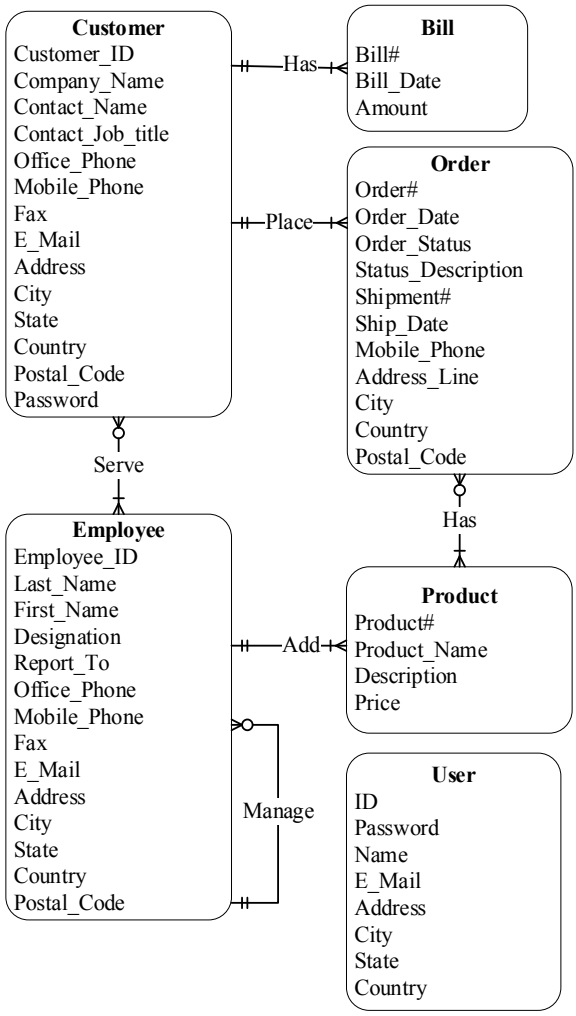


Fig. 4 Reference ERM

Customer*(Company_Name, Contact_Name, Contact_job_Title, Office_Phone, Mobile_Phone, Fax, EMail, Address, City, State, Country, Postal_Code)*
User*(ID, Password, Name, EMail, Address, city, state, country)*
Employee*(Employee_ID, Last_Name, First_Name, Designation, Office_Phone, Mobile_Phone, Fax, EMail, Address, City, State, Country, Postal_Code)*
Product*(Product#, Product_Name, Description, Price)*
Order*(order#, orderdate, Order_Status, Status_Description, Shipment#, Shipment_Date, Mobile_Phone, Adress, City, Country, Postal_Code)*
Bill*(Bill#, Bill_Date, Amount)*
Relationships
employee > 1,n (serve) customer > 0,n

Fig. 5 ERM for Validation



employee > 1,1 (manage) employee > 1,n
customer > 1,1 (place) order > 0,n
Customer > 1,1 (has) Bill > 1,n
order > 0,n (has) Product > 1,n

5.3 List of Errors and Suggestions

The output of this evaluation process is a list of errors with suggestions. The list of errors contains the sources along with the explanations.

- *customer—user* (Redundant Entities within the manual Model)
- *Address, order_status and shipment* (Missing or Mismatched Entities)
- *customer.address* (Convert Attributes *address* to Entity)
- *employee.address* (Convert Attributes *address* to Entity)
- *order.orderstatus* (Convert Attribute *orderstatus* to Entity)
- *order.orderstatuscode* (Move Attribute to *orderstatus*)
- *order.orderstatus* (Move Attribute to *orderstatus*)
- *order.statusdescription* (Move Attribute to *orderstatus*)
- *order.shipmentno* (Move Attribute to *shipment*)
- *order.shipmentdate* (Move Attribute to *shipment*)
- *product.unitinstock, product.reorderlevel, orderstatus.orderstatuscode, shipment.shipmentdescription* (Missing or Mismatched Attributes)
- *employee > 1,n(serve)customer > 0,n* (Wrong Maximum Cardinality of *Employee*)
- *employee > 1,1 (manage) employee > 1,n* (Wrong Minimum Cardinality of *Employee*)
- Missing Relationships

employee > 1,1(has)address > 1,n	*client > 1,1(has)address > 1,n*
shipment > 1,n(deliver)address > 1,1	*order > 1,1(has)invoice > 1,n*
order > 1,n(has)orderstatus > 1,1	*order > 1,1(posted)shipment > 1,1*

6 Performance Evaluation

We have applied mutation testing (by introducing faults intentionally in the targeted ERMs) for the evaluation of proposed approach. The performance is evaluated based on the number of errors detected from ERM. Results of the automated technique are compared with the manual assessment of models.

To differentiate the projected values and actual values, three classifications are being used i.e. True Positive (TP), False Positive (FP) and False Negative (FN). Recall (RCL) is to measure the correctly extracted elements by the proposed solution. Precision (PRC) is to measure the ratio between correct element generated by the approach and incorrect elements [41].

6.1 Statistics of ERMs Used for Validation

The following table contains statistics of models used in mutation testing for the validation of the proposed technique (Table 2).

Table 2 Statistics of models used for validation

Artifacts	Total
ERMs	63
Entities	6–38
Attributes	21–183
Relationships	5–36
Cardinalities	10–72

Table 3 Performance of structural analysis

	TP %	FP %	FN %	RCL %	PRC %
Entities	82.05	10.26	7.69	91.43	88.89
Attributes	82.61	9.57	7.82	91.35	89.62
Relationships	76.92	12.82	10.26	88.23	85.71
Cardinality	82.86	9.52	7.62	91.58	89.69

We used from 30 to 50% of the components for defect seeding.

6.2 Validation of Structural Analysis

We seeded the errors in the following manners for the validation of StrA.

- Copying and replacing the labels of the artifacts with the domain-based synonyms and for this process, we used some dictionary based tools.
- Moving the attributes to the irrelevant entities.
- Duplicating the attributes within and other entities.
- Swapping entities and attributes, entities and relationships.
- By deleting the components (Table 3).

6.3 Validation of Semantic Analysis

Following steps are performed to introduce errors in the labels for the validation of SA.

- Deleting and adding the characters.
- Adding special characters.
- Using the abbreviations.
- Changing the form of the verb (Table 4).

Table 4 Performance of semantic analysis

	TP %	FP %	FN %	RCL %	PRC %
Entities	80.65	12.90	6.45	92.59	86.21
Attributes	81.51	11.76	6.73	92.37	87.39
Relationships	82.05	10.26	7.69	91.43	88.89

7 Conclusion

In this effort, we presented an automated ERM assessment technique using the DK. The ERM extracted using our previous automated technique is used as a reference model for the evaluation of manually depicted model from the same set of requirements. We considered two important quality factors: completeness and correctness of the ERM. We applied StrA to find the sources of errors by evaluating the ERM. SA is to use DK for the matching of labels while performing StrA. The results in our case studies proved that it is an efficient technique. In the future, we are planning to extend this technique for the logical database design, physical database design, and other SE models.

References

1. Omar, N., Hanna, J.R.P., McKevitt, P.: Heuristic-based entity-relationship modelling through natural language processing. In: Artificial Intelligence and Cognitive Science Conference (AICS), pp. 302–313. Artificial Intelligence Association of Ireland (2004)
2. Meziane, F., Vadera, S.: Obtaining ER Diagrams Semi-Automatically From Natural Language Specifications (2004)
3. Btoush, E.S., Hammad, M.M.: Generating ER diagrams from requirement specifications based on natural language processing. Int. J. Database Theory Appl. **8**(2), 61–70 (2015)
4. Moody, D.L.: Theoretical and practical issues in evaluating the quality of conceptual models: current state and future directions. Data Knowl. Eng. **55**(3), 243–276 (2005)
5. Siau, K., Tan, X.: Improving the quality of conceptual modeling using cognitive mapping techniques. Data Knowl. Eng. **55**(3), 343–365 (2005)
6. Van Vliet, H., Van Vliet, H., Van Vliet, J.C.: Software Engineering: Principles and Practice, vol. 3. Wiley, New York (1993)
7. Stecklein, J.M., Dabney, J., Dick, B., Haskins, B., Lovell, R., Moroney, G.: Error Cost Escalation Through the Project Life Cycle (2004)
8. Boehm, B.W.: Software Engineering Economics, vol. 197. Englewood Cliffs (NJ), Prentice-hall (1981)
9. Stecklein, J.M., Dabney, J., Dick, B., Haskins, B., Lovell, R., Moroney, G.: Error Cost Escalation Through the Project Life Cycle (2004)
10. McGibbon, T.: Return on investment from software process improvement. DoD Softw. Tech. News **5**(4) (2002)
11. Pavlina, S.: Zero-defect software development. Dexterity Software (2003). www.dexterity.com
12. Rothman, J.: What does it cost to fix a defect?. Column Arch. Sticky Minds. Com (2002). http://www.stickyminds.com
13. Roman, G.C.: A taxonomy of current issues in requirements engineering. Computer **4**, 14–23 (1985)

14. Moody, D.L., Shanks, G.G.: Improving the quality of data models: empirical validation of a quality management framework. Inf. Syst. **28**(6), 619–650 (2003)
15. Moody, D.L.: Measuring the quality of data models: an empirical evaluation of the use of quality metrics in practice. In: ECIS 2003 Proceedings, p. 78 (2003)
16. Von Halle, B.: Data: asset or liability. Database Program. Des. **4**(7), 13–15 (1991)
17. Batini, C., Ceri, S., Navathe, S.B.: Conceptual Database Design: An Entity-Relationship Approach, vol. 116. Benjamin/Cummings, Redwood City, CA (1992)
18. Levitin, A., Redman, T.: Quality dimensions of a conceptual view. Inf. Process. Manage. **31**(1), 81–88 (1995)
19. Krogstie, J., Lindland, O.I., Sindre, G.: Towards a deeper understanding of quality in requirements engineering. In: Seminal Contributions to Information Systems Engineering, pp. 89–102. Springer, Berlin, Heidelberg (2013)
20. Nelson, H.J., Poels, G., Genero, M., Piattini, M.: A conceptual modeling quality framework. Softw. Qual. J. **20**(1), 201–228 (2012)
21. Hussain, T.: Revisiting quality metrics for conceptual models. In: TENCON 2014–2014 IEEE Region 10 Conference, pp. 1–6 (2014)
22. Javed, M., Lin, Y.: Iterative process for generating ER diagram from unrestricted requirements. In Proceedings of the 13th International Conference on Evaluation of Novel Approaches to Software Engineering, pp. 192–204 (2018). ISBN: 978-989-758-300-1
23. Krogstie, J.: Quality of Conceptual Data Models (2013)
24. Kesh, S.: Evaluating the quality of entity relationship models. Inf. Softw. Technol. **37**(12), 681–689 (1995)
25. Gosain, A., Nagpal, S., Sabharwal, S.: Quality metrics for conceptual models for data warehouse focusing on dimension hierarchies. ACM SIGSOFT Softw. Eng. Notes **36**(4), 1–5 (2011)
26. Thi, T.T.P., Helfert, M.: A Review of Quality Frameworks in Information Systems (2017). arXiv:1706.03030
27. Thomas, P., Smith, N., Waugh, K.: Automatically assessing diagrams. In: Proceedings of the IADIS International Conference on e-Learning (2009)
28. Batmaz, F., Hinde, C.J.: A Diagram Drawing Tool for Semi–Automatic Assessment of Conceptual Database Diagrams (2006)
29. Hussain, T., Awais, M.M., Shamail, S.: October. A fuzzy based approach to measure completeness of an entity-relationship model. In: International Conference on Conceptual Modeling, pp. 410–422. Springer, Berlin, Heidelberg (2005)
30. Lino, A.D.P., Rocha, A.: June. Automatic evaluation of ERD in e-learning environments. In: 13th Iberian Conference on Information Systems and Technologies (CISTI). IEEE (2018)
31. Choppella, V., Sengupta, A., Robertson, E.L., Johnson, S.D.: Constructing and Validating Entity-Relationship Data Models in the PVS Specification Language: A Case Study Using a Text-Book Example (2006)
32. Kazi, Z., Radulović, B., Berković, I., Kazi, L.: Ontology-based reasoning for entity-relationship data model semantic evaluation. Tehnički vjesnik **24(Supplement)**(1), 39–47 (2017)
33. Schenk, K.D., Vitalari, N.P., Davis, K.S.: Differences between novice and expert systems analysts: what do we know and what do we do? J. Manag. Inf. Syst. **15**(1), 9–50 (1998)
34. Wullink, M.: Data Model Maintainability A Comparative Study of Maintainability Metrics Improving (2010)
35. Moody, D.L.: Metrics for evaluating the quality of entity relationship models. In: International Conference on Conceptual Modeling. Springer, Berlin, Heidelberg, pp. 211–225 (1998)
36. Vachharajani, V., Pareek, J.: A proposed architecture for automated assessment of use case diagrams. Int. J. Comput. Appl. **108**(4) (2014)
37. Jayal, A., Shepperd, M.: The problem of labels in E-assessment of diagrams. J. Educ. Resour. Comput. (JERIC) **8**(4), 12 (2009)
38. Ruder, S., Ghaffari, P., Breslin, J.G.: Data Selection Strategies for Multi-domain Sentiment Analysis (2017). arXiv:1702.02426
39. Naili, M., Chaibi, A.H., Ghezala, H.H.B.: Comparative study of word embedding methods in topic segmentation. Procedia Comput. Sci. **112**, 340–349 (2017)

40. Mikolov, T., Chen, K., Corrado, G., Dean, J.: Efficient Estimation of Word Representations in Vector Space. Cornell University Library (2013). arXiv:1301.3781
41. Echeverría, J., Pérez, F., Pastor, Ó., Cetina, C.: Assessing the performance of automated model extraction rules. In: Advances in Information Systems Development, pp. 33–49. Springer, Cham (2018)

Analysis of Online Social Network after an Event

Myat Mon Oo and May Thu Lwin

Abstract Twitter can inform us what people are currently talking about and what is happening in the world. In the movement to free Reuters journalists (#FreeWaLoneKyawSoeOo) who were detained in Myanmar on Dec. 12, 2017, social media is utilized to express feeling and opinion. In social media, the rise of a trending topic leads to appear the temporal emergence of a set of users currently interested in that topic. Finding the influential person and measuring their influence are interesting problems in a social network. At first, the person who has many followers are more influence others. However, the influence score does not depend on the number of followers. In this study, we apply Social Network Analysis (SNA) to extract knowledge from social media data. We create a trending topic network graph related to an event. This graph was created by extracting relationships between users who are related to a given trending topic while they processed reply, mention or retweet operations. We used centrality measurement approaches and link analysis approach to find influential users. By using interaction relationships between users, we verify that PageRank can detect more influential users than other centrality measurement approaches. The result of this study will help to understand the structure of the network and also detect the most influential users.

Keywords Social network analysis · Trending topic · Network graph · Influential users · Social media

Myat Mon Oo (✉) · May Thu Lwin
University of Computer Studies, Mandalay, Myanmar
e-mail: myatmonoo@ucsm.edu.mm

May Thu Lwin
e-mail: maythulwin@ucsm.edu.mm

R. Lee (ed.), *Computer and Information Science*, Studies in Computational Intelligence 849, https://doi.org/10.1007/978-3-030-25213-7_11

1 Introduction

Social media has become an interesting and exciting area in which billions of individuals can share, interact, post and conduct information and daily activities via virtual communities and network. Facebook, Instagram, LinkedIn, Snapchat, Twitter, Viber, WeChat, Weibo, etc., are popular social media. The most popular social networks in the world that were ranked by the number of active users in January 2019, Facebook has 2,271,000,000 users, YouTube has 1,900,000,000 users and Twitter has 326,000,000 users, etc.[1] The enormous information on social media can attract data scientists, researchers, and companies who are trying to infer meaningful information.

Twitter is a popular social media in which user can create, share information in real time. Twitter (microblogging service) that enables users to share short text messages called tweets. Tweets were originally restricted to 140 characters, but on November 7, 2017, this limit was 280 characters for all languages except Chinese, Korean, and Japanese.[2] Twitter can be used for a variety of purposes which include URLs sharing, daily conversations and information news. By default, tweets are publicly visible, but a message can be restricted by senders delivery to just their followers. Tweets help us to know many stories on the web. Users may subscribe to the tweet of other users, known as "followings" and subscribers are known as "followers".

Users can forward others' tweets to their own feed, this process is known as a "retweet". The retweet mechanism is useful to spread of information or a new median of information dissemination, and can reach beyond of the original tweet's follower [1]. People often write a tweet in which a specific user is addressed, we call that user is a mention user. Mention indicates the ability of the user to communicate others in a conversation. Both mention and reply include" followed by the address of user's screen name. Users can also "like" (also known as "favorite") individual tweets.

In the social media such as Twitter, users can discuss various topic that they interested. So, some influential users can affect opinions, attitudes, behaviors or emotions of others [2]. The rise of a trending topic leads to appear the temporal emergence of a set of users currently interested in that topic [3]. A trending topic is a topic that are the dominant keywords or hashtags to describe a theme and can change every time depends on user interested.

In this paper, we analyzed an event about #FreeWaLoneKyawSoeOo. The movement to free Reuters journalists (#FreeWaLoneKyawSoeOo) has emerged on social media. #FreeWaLoneKyawSoeOo movement emerged as an interesting topic of a social movement on Twitter in Myanmar. Social network analysis (SNA) are applied in this study to extract the characteristics of trending topic network #FreeWaLoneKyawSoeOo on Twitter. Data used for this study were collected by crawling

[1] https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/.

[2] https://en.wikipedia.org/wiki/Twitter.

Twitter about above trending topic for a week, a social network has been created by 3001 users who include the hashtag #FreeWaLoneKyawSoeOo in their tweet, and 4293 social ties between them. And then we analyzed the structure for this trending topic graph.

Detection of the most influential users is an important task in social networks analysis. In recent years, the researchers interested to find the influential nodes in the network. There are many recent work to study influence and the diffusion of information in social networks. And in the network analysis, there are many works that has focused explicitly on sociometry which includes quantitative measures of influence, centrality or prestige and authority [4].

There are several methods based on the graph to find influential users. The centrality measurement methods such as degree centrality, betweenness centrality are essentially heuristics, usually intuitive notions are based such as to access and control over resources, or information brokerage. Finding the most influential user in a network is a conceptual problem. We do not have a definite method and measurement to identify social influence. Network measurement approaches are the main theories in the study of social networks to detect influential nodes. Centrality measure is the most empirical measurement in the social network. In practice, there were a large numbers of data in social network that can cause high time complexity and hard to measure centrality. Network analysis is very useful to identify important vertices within network.

In this paper, we focused on the task of identifying the most influential users in a social network working with trending topic network from services Twitter, that include relationships between users in the network. In order to do this, we began by collecting real and up-to-date data from the social networking platforms, namely, Twitter. In a trending topic social networks, different ranking algorithms were computed and the top-10 highest ranked users were extracted and analyzed. We contribute to characterize the structure of trending topic network and to identify the most influential users related to that trending topic based on relationships between users. We make the following objectives to achieve our aim:

(1) we created a trending topic network graph based on user interactions relations crawling data with a hashtag
(2) we analyze the characteristic of trending topic network
(3) we used network centrality measurement and link analysis approaches based on user relationships to identify influential users.

The remaining of this paper is organized as the following order. Section 2 presents related work and Sect. 3 discusses the methodology concerned with collecting data and finding influential users. In Sect. 4, we present result and discussion. Finally, Sect. 5 presents a conclusion of our work.

2 Related Work

In recent years, detection of topics over social media is dynamically development and receiving of user-generated contents are more and more interested in researchers [5]. There was enormous information on social media, researchers focus on emerging topic discovery in online content. Wang et al. [6] developed real-time earthquake detection system based on the number of words, the keywords in a tweet, etc, Schubert et al. [7] proposed "SigniTrend" emerging topic detection system with hashed significance thresholds. Mathioudakis and Koudas [8] proposed "TwitterMonitor" trend detection system using bursting keywords.

In the areas of communication theory and social science, the study of people's individual influence in a community has been an interesting topic since the 1940s [9]. Centrality measurement methods have been applied in the social network analysis to identify the most important nodes in the network. Many measurements of node centrality have been used commonly such as Betweenness centrality [10–12], Degree centrality [10], Closeness centrality [10] and so on. Catanese et al. [13] have applied betweeenness centrality in order to discover the central nodes of Facebook network.

Romero et al. [14] studied the degree measure to find the most influential users in Online Social Network (OSN). Liu et al. [15] proposed generalized closeness index (GCC) that maximize the distance among spreaders to identify multiple influential spreaders in complex networks. Liu et al. [16] proposed linear social influence modeling approach which focused on PageRank to measure important nodes in large-scale social networks. Chong et al. [17] studied how people adopt information about politic on Twitter using hashtag and how people facilitate to communicate among users with different political orientations.

In an online social network, the identification of influential users is an important task. Influence occurs when a person's emotion, opinions, or behaviors are affected by others. Watts et al. [18] defined the central idea of influence in marketing and diffusion research that is a minority of individuals who influence an exceptional number of their peers, and that are important to the formation of public opinion. Initially, a user who has numerous friends and followers is defined as influential users [19]. However, this concept recently evolved, and researchers commenced to utilize several related network structure and user's content to identify influential users [2, 20, 21].

Aftab et al. [22] detected influential nodes based on network metrics using social networks analysis. Sun et al. [20] intended to find influential users or leaders within a topic with follower/followee relationship between users. According to do this, firstly they measure the influence of online posts and then find the authors who wrote the most influential posts. Taniarza et al. [23] analyzed to determine the rank of users' influence score in a social media Twitter by adapting random-algorithm approach.

3 Methodology

Twitter provides to access the public information of tweet and users through the Twitter REST API v1.1. Twitter API is a good tool to investigate the profile information of non-protected users and the users who they follow and who they are followed by. However, Twitter limits the access rate and restricting the number of calls to maintain the reliability of its service and control costs. The rate limiting of the standard API is primarily on per-user-basic or per-user-access-token.

Currently, rate limits are divided into 15-minute intervals. The API version 1.1 limits an application based on the number of users and therefore granted it an access token. Each method in the API has it is own limit on the number of calls, however, if a method allows 15 calls per rate limit window. An application can make 15 calls to that method for each access token under its control.[3] Due to a limited query rate which has been imposed by network service providers and the large-scale nature of networks, crawling and enumerating content for each user account is computationally prohibitive [24]. NodeXL Basic is a Microsoft Excel plugin. The software can be used to obtain data from Twitter but it limits up to 2000 tweets. In this study, a dataset is created under a trending topic on Twitter for a week. We can programmatically crawl 9295 tweets using search API.

We used standard search API to collect the relevant tweets matching a specified query. This search API searches a sampling of tweets recently published in the past 7 days. All tweets that were returned by Twitter API are encoding with JavaScript Object Notation (JSON).[4]

We extract the edge list based on the relationships (reply or mention or retweet) between users. The resulting edge list is imported in NodeXL tool to analyze the structure of the network. We also used updated edge lists (removed duplicate edges to a single edge from NodeXL) and then applied it to draw the network graph and to detect influential users using NetworkX library, Python.

3.1 Using Graphs to Represent Social Relations

Social network analysis uses the social graph that consists of nodes (users) and edges (or ties) to represent the relationship between nodes. Graphs are usually used when representing networks, either undirected or directed. In this paper, we create a trending topic graph which consists of users and connections that are formed when people reply or mention or retweet one another concerned with keyword or hashtag. Figure 1 shows communication relationship types (reply, mention, retweet) between users.

[3] https://developer.twitter.com/en/docs.

[4] https://developer.twitter.com/en/docs/tweets/data-dictionary/overview/tweet-object.

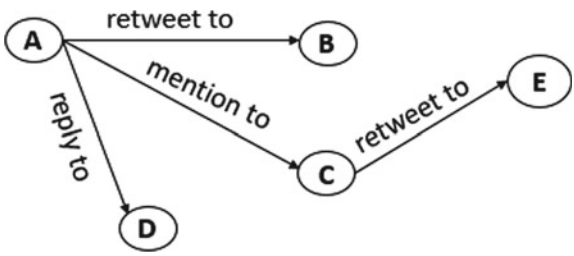


Fig. 1 Illustration of users and relationships between them

3.2 Data Collection

Analyzing social network about what is being occurred in the world and who is posting which messages in real time are really interesting. One of the methods which is commonly used to collect data from social media sites is application programming interfaces (APIs) . We crawled real-world data derived from a social media Twitter. The data related trending topic is a JSON file format, from which we created an edge list based on the relationship between users.

In this paper, the dataset was scraped from social media Twitter about the hashtag "#FreeWaLoneKyawSoeOo" from 2019-01-08 to 2019-01-14. We construct three different types of edges such as retweets, replies and mentions. In this study, a total of 3001 vertices generated a total of 4293 edges in #FreeWaLoneKyawSoeOo network. In this network, vertices represent the Twitter users while edges represent the communication relationships between the users.

3.3 Identification of Influential Users

To identify influential users we apply degree centrality [10, 25], closeness centrality [10], betweenness centrality [10–12, 26], and PageRank [27, 28].

Degree Centrality

Degree centrality defines as the counting of how many connections a node has. It considers the people have many connections is important. In an undirected graph, the number of edges of a node is the degree of that node. In an undirected graph, the degree centrality C_d of node v_i is

$$C_d(v_i) = d_i, \tag{1}$$

where d_i is the degree (number of adjacent edges) of node v_i.

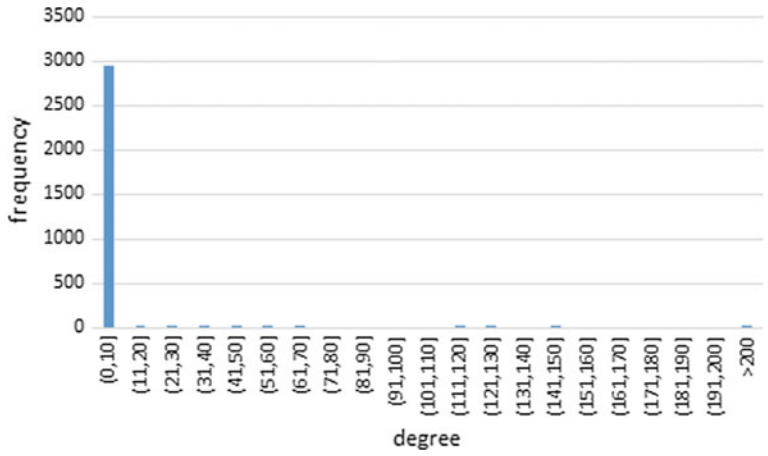


Fig. 2 Degree distribution

In a directed graph, the degree of the node is considered under three conditions, in-degree which counts the number of incoming links, out-degree which counts the number of outgoing links, and the total sum of in-degree and out-degree.

$$C_d(v_i) = d_i^{in} \qquad (prestige), \tag{2}$$

$$C_d(v_i) = d_i^{out} \qquad (gregariousness), \tag{3}$$

$$C_d(v_i) = d_i^{in} + d_i^{out}, \tag{4}$$

where d_i^{in} is the number of incoming links of node v_i and d_i^{out} is the number of outgoing links of of node v_i. Equation (4) is equivalent to Eq. (1), which calculates degree centrality of the undirected graphs.

We analyze three types of degree. In-degree, out-degree and degree distributions are highly right-skewed (Figs. 2, 3 and 4). This shows that there were so many nodes which have a small degree but a few nodes have a very large degree. The nodes with large-degree are often referred to as hubs. However, if we use degree distribution, we will capture only a small amount of the network structure because it ignores how the nodes are connected to each other. Top ten in-degree and out-degree users are listed in Table 1. Some of them are either news agencies or reporters, but consist of the individual users. And user#1407 and user#710 are consists of both degrees. So, these two users are both creators and distributors of information about that topic.

Betweenness Centrality

Betweenness centrality measures that which vertex lies on paths between other vertices. Vertices with high betweenness are the more influence vertices within a network that vertices can control information passing between others. Removing an impor-

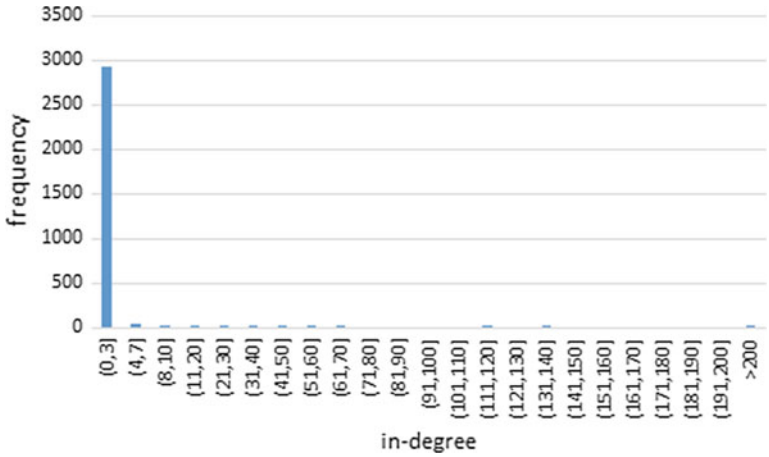


Fig. 3 In-degree distribution

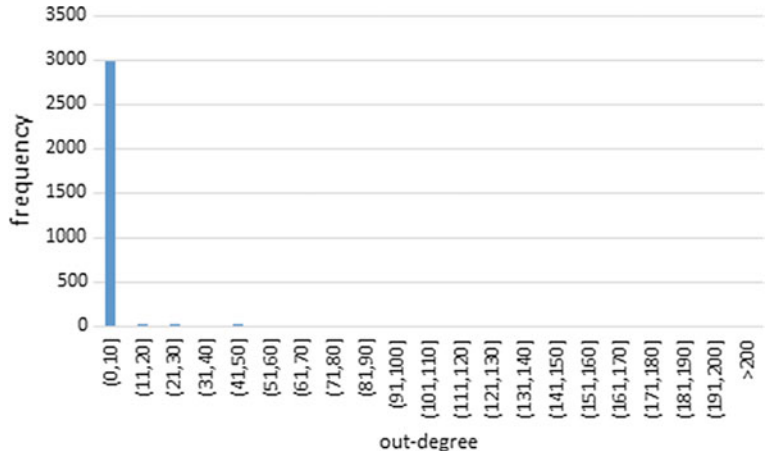


Fig. 4 Out-degree distribution

tant vertex from the network may cause disconnected communications between other vertices because it is located on the largest number of paths. Betweenness centrality [26] of a node v is the sum of the number of shortest paths that pass through v which is divided by the number of shortest paths from a source node to a target node:

$$C_B(v) = \sum_{s,t \in V} \frac{\sigma(s,t|v)}{\sigma(s,t)}, \tag{5}$$

Table 1 Top-10 degree, In-degree and Out-degree users

(a) Degree distribution

User	Degree
user#1435	1539
user#1157	542
user#1407	306
user#710	148
user#1048	122
user#846	115
user#1758	115
user#1163	65
user#234	64
user#2236	59

(b) In-degree distribution

User	In-degree
user#1435	1539
user#1157	542
user#1407	293
user#710	137
user#1048	120
user#846	114
user#1758	112
user#1163	63
user#234	59
user#2236	58

(c) Out-degree distribution

User	Out-degree
user#896	50
user#2754	22
user#1407	18
user#1202	16
user#815	14
user#2630	14
user#710	13
user#144	11
user#1251	11
user#2722	10

where $C_B(v)$ is the betweenness centrality of node v, V is the set of nodes, $\sigma(s, t)$ is the number of shortest paths between s and t, and $\sigma(s, t|v)$ is the number of those path passing through node v.

Closeness Centrality

Closeness centrality refers to the more central nodes are, the more quickly reach to the other nodes. Vertices with highest closeness centrality have more direct influence and accept information on other vertices. Closeness centrality [10] of a node u is the reciprocal of the sum of the shortest path distances from u to all $n - 1$ other nodes. Depends on the number of nodes in the graph, distances are summed and closeness is normalized by the sum of minimum possible distances $n - 1$. The formula of closeness centrality is defined as

$$C_C(u) = \frac{n-1}{\sum_{v=1}^{n-1} d(v, u)}, \tag{6}$$

where $C_C(u)$ is the closeness centrality of node u, $d(v, u)$ is the shortest-path distance between v and u, and n is the number of nodes in the graph.

PageRank

PageRank [28] refers to a node is important where if it linked from other important nodes and linked parsimonious nodes. Pagerank is a link analysis algorithm based on the concept of eigenvector centrality. Google use this algorithm to estimate the importance or popularity of a webpage, based on the in-link and out-link of that webpage [27]. The basic concept of PageRank is to propagate of influence along the network of web pages, instead of counting the number of inlink of a web page. The formula of PageRank algorithm is

$$PR(u) = \frac{1-d}{N} + d \sum_{v \in B(u)} \frac{PR(v)}{N_v}, \tag{7}$$

where $PR(u)$ is a PageRank score of a page u, $B(u)$ is the set of pages that point to u, $PR(u)$ and $PR(v)$ are rank scores of page u and v, respectively. N is the number of web pages, N_v denotes the number of outgoing links of page v and d is a damping factor that is usually set to 0.85.

4 Result and Discussion

In this paper, Python tool is used. Python is a very powerful programming language, which is created by Guido van Rossum and it is easy to learn. We used networkX library with spring layout and Pyplot function to create a directed graph. NetworkX is used for visualization of the network.

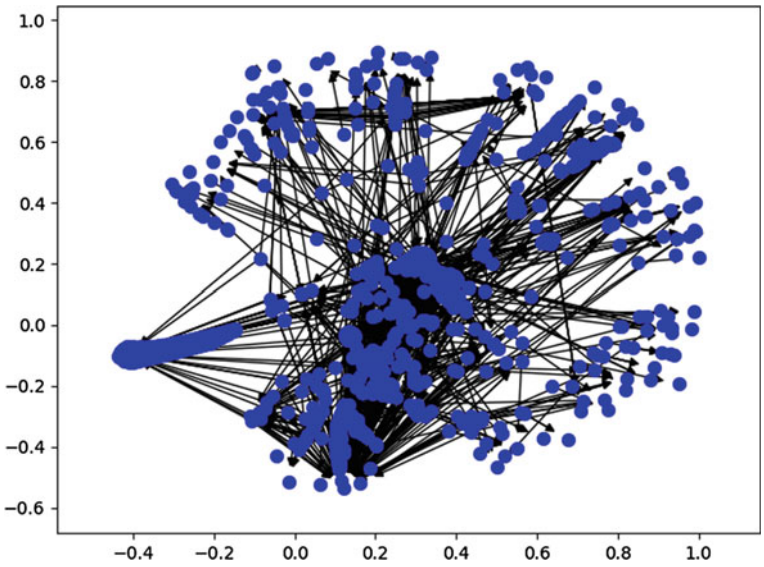


Fig. 5 A trending topic network graph based on the retweet, mention and reply relationship

Figure 5 shows the network of a trending topic #FreeWaLoneKyawSoeOo. We used NodeXL tool to analyze the structure of the network (Table 2). NodeXL is an Excel template and it has a lot of features such as networking visualization, social network analysis, advanced network metrics and so on.

4.1 Structure and Characteristic of Network Graph

Graph theory is used to analyze the structure of the social network in SNA. In Table 2, the graph is a directed graph, the vertices show 3001 users that are engaged in a microblogging conversation including #FreeWaLoneKyawSoeOo as a topic area. The number of edges is 4293, which suggests quite a bit of connectivity (retweet, mentions, and replies) between these vertices or nodes.

Edges include repeated vertices pairs, in this study 4293 edges consists of duplicated 393 edges. When user u replies to user v many times, duplicate vertices pairs can occur. These pairs can cause some metrics, such as *degree*, to be inaccurate. So, the 393 edges were converted into a single weighted edge.

The graph does not contain self-loops which is an edge that connects a vertex to itself. Self-loops effect in-degree and out-degree of the graph. The average geodesic distance in this network is 3.11 hops or steps, which means that the members of this social network are in fairly close proximity in terms of reaching each other.

Table 2 Overall metrics

Graph Metric	Value
Graph type	Directed
Vertices	3001
Unique edges	3900
Edges with duplicates	393
Total Edges	4293
Self-Loops	0
Reciprocated vertex pair ratio	0.002221674
Reciprocated edge ratio	0.004433498
Connected components	27
Single-Vertex connected components	0
Maximum vertices in a connected component	2929
Maximum edges in a connected component	4241
Maximum geodesic distance (Diameter)	9
Average geodesic distance	3.109094
Graph density	0.000450961

The maximum geodesic distance (diameter) is 9, or 9 hops separating the two most distance nodes in this network.

Table 3 shows the top 10 ranking result lists of betweenness, closeness and PageRank centralities that have been calculated by using NetworkX library. We apply PageRank to the trending topic network. The top 10 list that was calculated by PageRank has the same user as in-degree list except for user#481 and user#2807. The two users namely user#1163 and user#2236 are dropped from the in-degree list and some rank has changed. Although the two ranked lists do not match exactly, users are ranked similarity by the number of in-degree and PageRank.

By comparing the ranking result of Closeness and PageRank, seven users are the same with a little changed of rank. According to Table 3, there were 18 unique users from the lists of the top 10 with betweenness, closeness, and PageRank centralities. The rates of common users in three ranking lists are 0.28, 0.39 and 0.50, respectively. It can be concluded that the top 10 ranked users categorized by the PageRank are more vital than the other two measurements in the #FreeWaLoneKyawSoeOo network. In top-10 PageRank list, top-1 is user#1435 who has 1539 in-degree and no out-degree. So, we conclude that his post is shared by 1539 users and he didn't share other's post. He has high prestige and his description on his home page is "Vice President Mike Pence. Husband, father, and honored to serve as the 48th Vice President of the United States.". Top-2 is user#1157 who has 1539 in-degree and no out-degree. We conclude that this account is Reuters, a news agency so this account is a source of

Table 3 TOP-10 users with betweenness, closeness and pagerank centrality

Rank/Centrality	Betweenness	Closeness	PageRank
Top 1	user#1407	user#1435	user#1435
Top 2	user#710	user#1157	user#1157
Top 3	user#2754	user#1407	user#1407
Top 4	user#1758	user#710	user#710
Top 5	user#846	user#481	user#1048
Top 6	user#234	user#1048	user#846
Top 7	user#2231	user#2556	user#481
Top 8	user#1119	user#1427	user#1758
Top 9	user#711	user#2273	user#2807
Top 10	user#1287	user#1758	user#234

information. Top-3 is user#1407 who has 239 in-degree and 18 out-degree. He is a Reuters journalist so, he is not only a creator of the information but also a distributor of that information in the #FreeWaLoneKyawSoeOo network.

In this work, the data set is used during a week (from 2019-01-08 to 2019-01-14) scrapped from Twitter. The result of this study will help to understand the structure of a trending topic graph and detect the most influential users.

5 Conclusion

Social media is a medium by using it people can create, share and access the data through networks. In this paper, we create a trending topic network graph scrapping the tweets related to an event. The trending topic graph was created by collecting of connections that are formed when people reply or mention or retweet one another concerned with keyword or hashtag. The relationship between users who are related to a given trending topic can be detected. The result of this study will help the structure of a trending topic graph and the most influential users can also be detected. Network measurement methods (Betweenness Centrality, Closeness Centrality, and PageRank) which can identify the most influential users are applied. In this identification process, PageRank algorithm can detect the most influential person than Betweenness and Closeness Centralities Algorithms.

References

1. Kwak, H., Lee, C., Park, H., Moon, S.: What is Twitter, a social network or a news media? In Proceedings of the 19th International Conference on World Wide Web, pp. 591–600. ACM, Apr 2010
2. Jianqiang, Z., Xiaolin, G., Feng, T.: A new method of identifying influential users in the microblog networks. In: IEEE Access, vol. 5, no. c, pp. 3008–3015 (2017)

3. Recalde, L., Nettleton, D.F., Baeza-Yates, R., Boratto, L.: Detection of Trending Topic Communities, pp. 205–213 (2017)
4. Subbian, K., Melville, P.: Supervised rank aggregation for predicting influencers in twitter. In Proceedings of 2011 IEEE International Conference on Privacy, Security, Risk and Trust IEEE International Conference on Social Computing PASSAT/SocialCom 2011, pp. 661–665 (2011)
5. Miao, Z., et al.: Cost-Effective online trending topic detection and popularity prediction in microblogging. ACM Trans. Inf. Syst. **35**(3), 1–36 (2016)
6. Wang, X., Zhu, F., Jiang, J., Li, S.: Real time event detection in twitter. Lecture Notes in Computer Science (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), LNCS, vol. 7923, pp. 502–513 (2013)
7. Schubert, E., Weiler, M., Kriegel, H.-P.: SigniTrend: Scalable detection of emerging topics in textual streams by hashed significance thresholds. In: Proceedings of 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining—KDD '14, pp. 871–880 (2014)
8. Mathioudakis, M., Koudas, N.: Twittermonitor: trend detection over the twitter stream. In: Proceedings of 2010 ACM SIGMOD International Conference on Management of Data, pp. 1155–1158 (2010)
9. Ma, X., Li, C., Bailey, J., Wijewickrema, S.: Finding Influentials in Twitter: A Temporal Influence Ranking Model (2017)
10. Freeman, Linton C.: Centrality in social networks conceptual clarification. Soc. Netw. **179**(1968), 215–239 (1978)
11. Newman, M.E.J.: A measure of betweenness centrality based on random walks. Soc. Netw. **27**(1), 39–54 (2005)
12. Brandes, U.: A faster algorithm for betweenness centrality. J. Math. Sociol. **25**(2), 163–177 (2001)
13. Catanese, S., De Meo, P., Ferrara, E., Fiumara, G., Provetti, A.: Extraction and analysis of facebook friendship relations. In: Computational Social Networks, pp. 291–324. Springer, London (2012)
14. Romero, D.M., Asur, S., Galuba, W., Huberman, B.A.: Influence and Passivity in Social Media Background on Twitter, pp. 18–33. Springer, Berlin, Heidelberg (2011)
15. Liu, H.L., Ma, C., Xiang, B.B., Tang, M., Zhang, H.F.: Identifying multiple influential spreaders based on generalized closeness centrality. Phys. A Stat. Mech. Appl. **492**, 2237–2248 (2018)
16. Liu, Q.I., et al.: An Influence Propagation View of PageRank, vol. 11, no. 3, (2017)
17. Chong, M.: Analyzing political information network of the U.S. Partisan public on twitter. Lecture Notes in Computer Science (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), LNCS, vol. 10766, pp. 453–463 (2018)
18. Watts, D.J., Dodds, P.S.: Influentials, networks, and public opinion formation. J. Consum. Res. **34**(4), 441–458 (2007)
19. Cha, M., Haddai, H., Benevenuto, F., Gummadi, K.P.: Measuring user influence in Twitter: The million follower fallacy. In: International AAAI Conference on Weblogs and Social Media, pp. 10–17 (2010)
20. Sun, B., Ng, V.T.Y.: Identifying influential users by their postings in social networks. In: Lecture Notes in Computer Science (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), LNAI, vol. 8329, pp. 128–151 (2013)
21. Farahani, H.S., Bagheri, A., Saraf, E.H.K.M.: Characterizing behavior of topical authorities in Twitter. In: Proceedings of IEEE International Conference on Innovative Mechanisms for Industry Application ICIMIA 2017, no. Icimia, pp. 581–586 (2017)
22. Farooq, A., Joyia, G.J., Uzair, M., Akram, U.: Detection of influential nodes using social networks analysis based on network metrics. In: Proceedings of 2018 International Conference on Computing, Mathematics and Engineering Technologies Inven. Innov. Integr. Socioecon. Dev. iCoMET 2018, vol. 2018, pp. 1–6, Jan 2018
23. Taniarza, N., Adiwijaya, Maharani, W.: Social network analysis using k-Path centrality method. J. Phys. Conf. Ser. **971**(1) (2018)

24. Wang, P., Zhao, J., Lui, J.C.S., Towsley, D., Guan, X.: Fast crawling methods of exploring content distributed over large graphs. Knowl. Inf. Syst. 1–26 (2018)
25. Zafarani, R., Abbasi, M.A., Liu, H.: Social Media Mining: an Introduction. Cambridge University Press (2014)
26. Brandes, U.: On variants of shortest-path betweenness centrality and their generic computation. Soc. Netw. **30**(2), 136–145 (2008)
27. Langville, A.N., Meyer, C.D.: Google's PageRank and Beyond: The Science of Search Engine Ratings (Book Review). Springer Science Media **2** (2008)
28. Brin, S., Motwani, R., Page, L., Winograd, T.: What can you do with a web in your pocket? IEEE Data Eng. Bull. **21**(2), 37–47 (1998)

Hybrid Radius Particle Swarm Optimization Applications

Mudarmeen Munlin

Abstract The Hybrid Radius Particle Swarm Optimization (HRPSO) algorithm is developed, where the social interaction among the agent particle over the radius of swarm circle topology is applied to avoid premature convergence. For the optimization benchmark functions, HRPSO performs better than Particle Swarm Optimization (PSO) and other existing methods. The HRPSO is applied to solve several real-world optimization problems, including the Resource-Constrained Project Scheduling Problem (RCPSP) and Travelling Salesman Problem (TSP). We have designed and investigated different approaches, such as adaptive mutation, forward backward propagation, and k-means combined with the Radius Particle Swarm Optimization (RPSO) to solve these problems. The efficiency of the proposed method is tested against the existing methods. The results show that the HRPSO gives better optimum results.

Keywords Resource constrained project scheduling problem · Particle swarm optimization · Radius particle swarm optimization · Adaptive mutation · Forward-backward improvement · K-means

1 Introduction

Optimization problems have been studied over the past few decades, with many different techniques and algorithms being adapted and used. Usually, optimization methods can be separated into two different groups: exact and approximate algorithms [1]. Exact algorithms guarantee that the obtained result is the optimum solution. In some extreme examples, exact algorithms, such as the Branch and Bound algorithm and Dynamic Programming, are known to be time intensive. This is the reason why most previous investigations use exact algorithms for small numbers of instances in optimization problems. When the number of instances becomes large,

M. Munlin (✉)
Faculty of Information Science and Technology, Mahanakorn University of Technology, Bangkok 10530, Thailand
e-mail: mmunlin@gmail.com

R. Lee (ed.), *Computer and Information Science*, Studies in Computational Intelligence 849, https://doi.org/10.1007/978-3-030-25213-7_12

approximate algorithms, such as heuristic and metaheuristic algorithms, are often used. The advantages of the metaheuristic algorithms include computational feasibility and effectiveness with acceptable time and memory requirements [2]. Unlike exact algorithms metaheuristic algorithms may not yield a best solution; however, they can obtain a sufficiently good solution within a reasonable amount of time.

A metaheuristic algorithm is a stochastic algorithm that involves randomization and local searching. The randomization process generates solutions that explore the search space and are responsible for reaching a global solution. The local search is used to determine convergence and focus on achieving a good solution in a specific region [2]. Most metaheuristic techniques draw inspiration from nature, such as Genetic Algorithm, Ant Colony Optimization and Particle Swarm Optimization (PSO). These algorithms use the population or the swarm to candidate the solution. However, the metaheuristic algorithm suffers from the premature convergence problem when facing a complex optimization problem. Premature convergence may also be due to trapping of a population in the local minima while searching for a solution to complex problems. These trapped populations will not participate in further search, leading to premature convergence and poor-quality outcomes.

As a result, the metaheuristic algorithm will always be successful on a given optimization problem if it can furnish a balance between global exploration and the exploitation. The global exploration generally refers to the ability to visit many different regions of the search space, whereas the local exploitation refers to the ability to obtain high quality solutions within those regions. Thus, it is very important to select a metaheuristic algorithm for each optimization problem.

The resource-constrained project scheduling problem (RCPSP) is a classical problem in the group of scheduling and NP-Hard. The RCPSP is a complex problem which involves resources and precedence constraints. Many studies have solved the RCPSP by applying the methods and the mechanisms to find the better make-span and minimize the time of the scheduling. These methods include exact, heuristic and metaheuristic procedures. Many papers solve the RCPSP by applying the meta-heuristics and the heuristics-based schemes, such as Bee Algorithm, shuffled frog-leaping algorithm (SFLA), hybrid ant colony optimization (ACO), artificial immune algorithm (AIA), scatter search algorithm (SS), hybrid genetic algorithm (GA) [3], the particle swarm optimization algorithm (PSO) [4]. These methods are tested against the standard dataset of the project scheduling problem library (PSPLIB).

The Traveling Salesman Problem (TSP) is a well-known combinatorial optimization problem [5]. More specifically, it is used as a benchmark for many optimization research techniques owing to its computational complexity. The TSP has been extensively studied by the researchers interested in the combinatorial optimization problem and application to real-world including network routing, industry, logistic, and transportation. Current trends in decision management are required for cost reduction and delivery. Given the distance between all pairs of cities, the goal of TSP is to find an efficient tour to all cities. The TSP is the mathematical problem to find the minimum length of a tour. The difficulty of the problem has been recognized and studied for many years. Several different methods have been proposed to solve this problem. The methods include genetic algorithm, ant colony optimization, particle

swarm optimization and honey bees mating optimization algorithm. Due to the TSP is NP-hard, only small instances can be solved to optimality.

We propose a Radius Particle Swarm Optimization (RPSO) algorithm to locate the agent particle within the radius of a circle. A radius neighborhood is defined for each individual particle as the subset of particles. Each particle in the overlap radius can be in multiple groups. Once, a group is defined, we locate the best particle in that group and assign to the agent particle. The agent particle is the candidate for finding the optimal solution. The radius parameter is calculated based on some benchmark functions. We hybridize the K-means and RPSO to solve the traveling salesman problem, employ adaptive mutation and the forward-backward improvement (FBI) method to improve the quality of the resulting schedule of the RCPSP. The performance of hybrid-RPSO (HRPSO) algorithm is tested against the traditional PSO, and others existing methods.

2 HRPSO for the Resource-Constrained Project Scheduling Problem

The Hybrid Radius Particle Swarm Optimization (HRPSO) is proposed for the Resource-Constrained Project Scheduling Problem (RCPSP) with the minimum makespan as the objective function. It is then applied to incorporate adaptive mutation to solve the impact on the bounds of search space so that the diversity among particles gradationally decreases. In addition, we introduce an integrated method based on hybridization of HRPSO and a local search algorithm to solve the RCPSP. The general framework of the HRPSO for RCPSP is illustrated in Fig. 1. In the beginning, we randomly generate the initial position value in each dimension using a priority-based solution. Next, the optimization loop starts with the adaptive mutation to solve the impact on the bounds of search space. Afterwards, the fitness of the particle is evaluated by mapping the position of particle with the Priority-Serial Schedule Generation Schema (P-SSGS). Forward-Backward Improvement (FBI) is then applied to the current solution to improve convergence speed and reduce the makespan of the schedule. Eventually, the HRPSO takes advantage of RPSO to maintain the swarm diversity and evolution by sharing information from the agent particles, which effectively keeps the balance between the global exploration and the local exploitation. The example shown in Fig. 2 is a project with 11 activities; the particle will fly in the search space with 11 dimensions. Each dimension of particle represents an activity and its corresponding value presents the priority in the range of [0.0, 1.0], the higher value represents the higher priority.

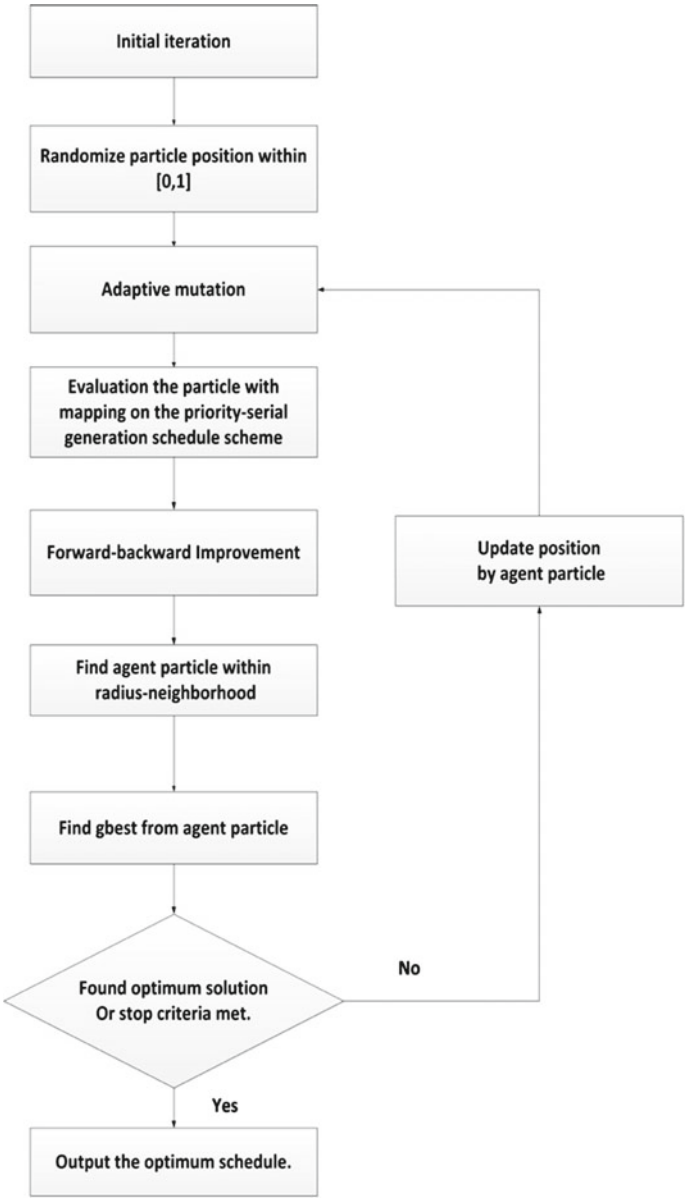


Fig. 1 The general framework of the HRPSO for RCPSP

j_0	j_1	j_2	j_3	j_4	j_5	j_6	j_7	j_8	j_9	j_{10}
A1	*A2*	*A3*	*A4*	*A5*	*A6*	*A7*	*A8*	*A9*	*A10*	*A11*
0.9	0.85	0.7	0.8	0.75	0.65	0.6	0.55	0.5	0.45	0.4

Fig. 2 Representation of priority in each dimension

2.1 Adaptive Mutation

According to the PSO algorithm, each particle is limited by the bounds of the search area (v_{max} is the upper bound and v_{min} is the lower bound) to control the flying area. Consequently, the velocity is effective at searching for the regional solution, such that the velocity may be larger than v_{max} or smaller than v_{min} and unable to escape from the bounds of the search area. We investigate here, according to RPSO algorithm, which extends PSO represents lower and upper bounds that are usually set between 0.0 and 1.0 for the search space of the RCPSP. When the velocity is larger than the upper bound (1.0) or smaller than the lower bound (0.0), a mutation operator will be triggered (Fig. 3) in order to escape from the bounds of the search area. Adaptive mutation is given in Algorithm 1.

Algorithm 1. Adaptive Mutation

Input Particle Position

Output New Particle Position

1 Randomly $x'_{ij}(t)$

2 **Select** the mutation points, $\varepsilon_1, \ldots, \varepsilon_n \in [1, m] \forall\, x_{ij}(t+1)[0.0,1.0]$

3 **for** $j = 1, \ldots, \varepsilon_n$ **do**

4 **if** Random [0,1] $> P_{mut}$ **then** $x_{ij}(t+1) \leftarrow random[0,1]$

else $x_{ij}(t+1) \leftarrow x'_{ij}(t)$

5 **end for**

6 **Return** $x_{ij}(t+1)$

			ε_1		ε_2	ε_n					
$x_{ij}(t+1)$	0.22	0.21	1.0	0.40	0.0	1.0	0.09	0.59	0.04	0.77	0.37
$x'_{ij}(t)$	0.67	0.34	0.45	0.18	0.01	0.53	0.96	0.81	0.70	0.30	0.54
$x_{ij}(t+1)$	0.22	0.21	0.45	0.40	0.01	0.53	0.09	0.59	0.04	0.77	0.37

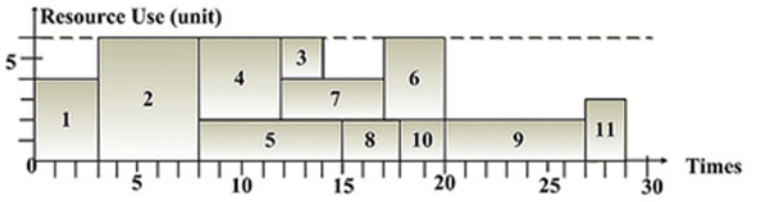
Fig. 3 Adaptive mutation

2.2 *Forward-Backward Improvement*

Valls [6] proposed the Forward-Backward Improvement (FBI). There are two processes involved in the FBI. Firstly, in the backward process, the activities are arranged from right to left and scheduled at the last finish time (LFT) are shifted to the right. After that, in the forward process, the activities are arranged from left to right and scheduled at the earliest start time (EST) are shifted back to the left.

Figure 4a–e illustrates a single-iteration step generated by FBI. In order to reduce the makespan of schedule in Fig. 4a, the backward procedure is employed by descending order from the LFT (Fig. 4b) and shifting the activities to the right as much as possible. Activity number 11, 9 and 6 cannot be scheduled later. Activity number 10, 8 and 7 can be shifted rightwards to a new LFT. Shifting activity numbers 10, 8 and 7 makes additional resources available and thus the LFT of the later activities can be adjusted (Fig. 4c). The backward procedure is performed in ascending order of the EST (Fig. 4d) and further improvement of the schedule is possible by shifting as many of the activities to the left as possible. For the example shown Fig. 4e, the FBI procedure reduces the total makespan by five time units. Obviously, the FBI can improve the quality of the schedule and simultaneously guarantee the precedence feasibility of the generated schedule.

Upon conclusion of each iteration of the HRPSO framework, a subset of particles contains the minimum makespan or the agent particle from the search space of the swarm circle topology. The HRPSO finds the agent particle by regrouping the particles within a given radius and determines the agent particle, which is the best particle of the group, for each local optimum. The procedure iterates by updating the new particle position using its agent particle. The HRPSO can maintain appropriate swarm diversity and jump out of local optima using the agent particle to achieve the global optimum. When the global optimum is achieved or the procedure reaches the maximum number of iterations, the schedule with minimum makespan obtained by the HRPSO is returned as the output.

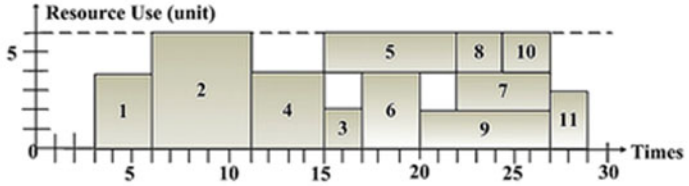


Activity List	1	2	5	4	8	7	3	10	6	9	11
EST	1	3	8	8	15	12	12	18	17	20	27
LFT	3	8	15	12	18	17	14	20	20	27	29

(a)

Activity List	1	2	4	3	5	7	8	10	6	9	11
EST	1	3	8	12	8	12	15	18	17	20	27
LFT	**3**	**8**	**12**	**14**	**15**	**17**	**18**	**20**	**20**	**27**	**29**

(b)

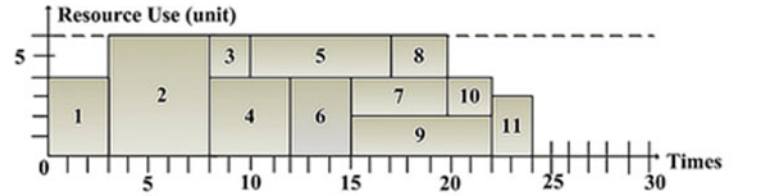


Activity List	1	2	4	3	5	7	8	10	6	9	11
EST	3	6	11	15	15	22	22	25	17	20	27
LFT	6	11	15	17	22	27	25	27	20	27	29

(c)

Activity List	1	2	4	3	5	6	9	7	8	10	11
EST	**3**	**6**	**11**	**15**	**15**	**17**	**20**	**22**	**22**	**25**	**27**
LFT	6	11	15	17	22	20	27	27	25	27	29

(d)



Activity List	1	2	4	3	5	6	9	7	8	10	11
EST	1	3	8	8	10	12	15	15	17	20	22
LFT	3	8	12	10	17	15	22	20	20	22	24

(e)

Fig. 4 The forward-backward improvement procedure

2.3 Experiments and Results

The experiment employs the well-known dataset called PSPLIB. This dataset contains subset J30, J60, and J120 which have 30, 60, and 120 activities respectively. All the dataset instances are designed with different network complexity, resource factor and resource strength. The set J30 and J60 consists of 480 instances (48 groups), and J120 consists of 600 instances (60 groups), which each group includes 10 benchmark instances. Hence, there are total of 1560 instances. The PSPLIB dataset has 4 renewable resources and each activity has 3 successors. The dataset is available at http://www.om-db.wi.tum.de/psplib/data.html with the best-known solution values. Only the set J30 are the optimal values for all instances known. In the set J60 and the set J120 the optimal values for some instances are not known and only upper bounds and lower bounds are provided. The upper bound is the current best solutions.

The acceleration constants c_1, c_2 are both *1.42694*, *Pmut* equals 0.5, x_{max} and v_{max} are set to equal and in the range of [0.0, 1.0]. The parameter w used a linear decreasing, which change from 0.9 to 0.4. For the HRPSO, μ is 0.4. The termination condition for each instance is the limit of the number generated schedules and is chosen at 1000, 5000 and 50,000. The number of generated schedules can be calculated by Eq. (1).

$$Generated\ schedules = Populationsize \times Iteration\ number \tag{1}$$

The number of generated schedules is 960 (60 particles × 16 iterations) for a limit of 1000 schedules, 4980 (60 particles × 83 iterations) for a limit of 5000 schedules and 49,980 (60 particles × 833 iterations) for a limit of 50,000 schedules. We run the H-RPSO 30 times independently for each instance and set the maximum number of schedules to 1000, 5000 and 50,000 as the stop criterion. The average response variable (*ARV*) values for J30, J60, and J120 are the average deviation values for the total instances (B). B is 480 and 600 respectively. The *ARV* and the optimum rate are given in Eqs. (2) and (3) respectively.

$$ARV = \sum_{l=1}^{B} \sum_{k=1}^{30} \frac{(OPT - makespan_{k,l})}{OPT} / 30 \cdot B \tag{2}$$

$$Optimum\ Rate = \frac{Optimum\ Instance}{Total\ Instance} \times 100 \tag{3}$$

where $makespan_{kl}$ is the make-span of the lth instance obtained by the HRPSO in the kth and *OPT* is the lower bound optimum. For all dataset, the HRPSO performs best with 1000, 5000 and 50,000 schedules, i.e. for J30 ARV is 0.11% with 1000 schedules, 0.03% with 5000 schedules and 0.01% with 50,000 schedules, and optimum rate is 94.17, 97.50 and 98.33% respectively. Thus, the make-span obtained by HRPSO is much closer to the optimal solution than the classic PSO. In order to compare the performance of the HRPSO with that of the best algorithms published so far, we have solved the dataset of instances with the HRPSO using 1000,

5000 and 50,000 generated schedules. Results are drawn in Tables 1, 2 and 3 with the average response variable and optimum rate. We can see that the HRPSO outperforms the metaheuristic methods previously published in the literature under the same schedules.

Table 1 Average response variable and optimum rate for J30

Algorithm	ARV			Optimum rate (%)		
	1000	5000	50,000	1000	5000	50,000
HRPSO	**0.11**	**0.03**	**0.01**	**94.17**	**97.50**	**98.33**
PABC [7]	0.34	0.17	–	86.60	91.74	–
BSO [8]	0.65	0.36	0.17	77.30	85.63	92.09
BA [8]	0.63	0.33	0.16	78.54	86.25	92.50
GAPS [9]	0.06	0.02	0.01	–	–	–
Hybrid-GA [10]	0.27	0.06	0.02	–	–	–

Table 2 Average response variable and optimum rate for J60

Algorithm	ARV			Optimum rate (%)		
	1000	5000	50,000	1000	5000	50,000
HRPSO	**11.54**	**10.23**	**10.11**	**75.41**	**78.33**	**80.20**
PABC [7]	12.35	11.96	–	72.5	74.03	–
BSO [8]	13.67	12.7	12.45	64.38	70.63	71.88
BA [8]	13.35	12.83	12.41	66.25	68.34	71.67
GAPS [9]	11.72	11.04	10.67	–	–	–
Hybrid-GA [10]	11.56	11.1	10.73	–	–	–

Table 3 Average response variable and optimum rate for J120

Algorithm	ARV			Optimum rate (%)		
	1000	5000	50,000	1000	5000	50,000
HRPSO	**34.50**	**31.94**	**30.25**	**30.16**	**33.50**	**35.83**
PABC [7]	36.84	35.79	–	29.5	31.2	–
BSO [8]	41.18	37.86	35.7	17	22.5	25.17
BA [8]	40.38	38.12	36.12	17.84	20.84	24.5
GAPS [9]	35.87	33.03	31.44	–	–	–
Hybrid-GA [10]	34.07	32.54	31.24	–	–	–

3 HRPSO for the Travelling Salesman Problem

The difficulty of the Travelling Salesman Problem (TSP) has been recognised for many years; however, the TSP is NP-hard and it may be challenging to design the exact algorithm to produce the optimal solution in a reasonable time. Only small instances of TSP can be solved optimally. We employ k-means and Radius Particle Swarm Optimization to solve the TSP with the shortest tour using the divide-and-conquer strategy. In the main, a solution to the TSP is decoded the tour from the position with cities clustering. Therefore, cities clusters are detached by the k-means algorithm and then ordered in a sequence of sub-cities and consolidated using RPSO. Thus, the data clustering allows for the improvement of the algorithm to reduce the complexity of the TSP problem. For solving the TSP, we construct an integrated framework based on hybridization of k-means and RPSO and the local search algorithm to resolve the crossing and jagged path problems as shown in Fig. 5.

3.1 k-Means

At the beginning, we use the k-means algorithm to create city clusters. Afterwards, we determine the position value in each dimension of the RPSO by a priority-based solution and then appoint the parameter of the RPSO. Next, the optimization loop starts with the adaptive mutation to solve the impact on the bounds of search space. Furthermore, the fitness of the particle is evaluated by decoding the position of a particle with city clusters. Concretely, the local search, with 2-OPT, is then applied to the current solution to reduce the crossing and jagged path problems and improve convergence speed.

An example of the distribution of the TSP instance is shown in Fig. 6, we divide cities into four clockwise quadrants so that the sum of the mean squares of the distances between nodes is minimum. In addition, we construct a new feature vector for each city in which the distance (z) between the geographical centre and the coordinate of cities is calculated with Eqs. (4), (5) and (6).

$$\delta_x = \frac{1}{n}\sum_{i=1}^{n} \rho_{x_i} \tag{4}$$

$$\delta_y = \frac{1}{n}\sum_{i=1}^{n} \rho_{y_i} \tag{5}$$

$$z_i = \sqrt{\left(\delta_x - \rho_{x_i}\right)^2 + \left(\delta_y - \rho_{y_i}\right)^2} \tag{6}$$

where (δ_x, δ_y) represents the geographical center, (ρ_{x_i}, ρ_{y_i}) represents the coordinate of the ith city and n represents the number of cities. The city clustering using k-means algorithm is elaborated as follows: To begin we define how many clusters (k) are

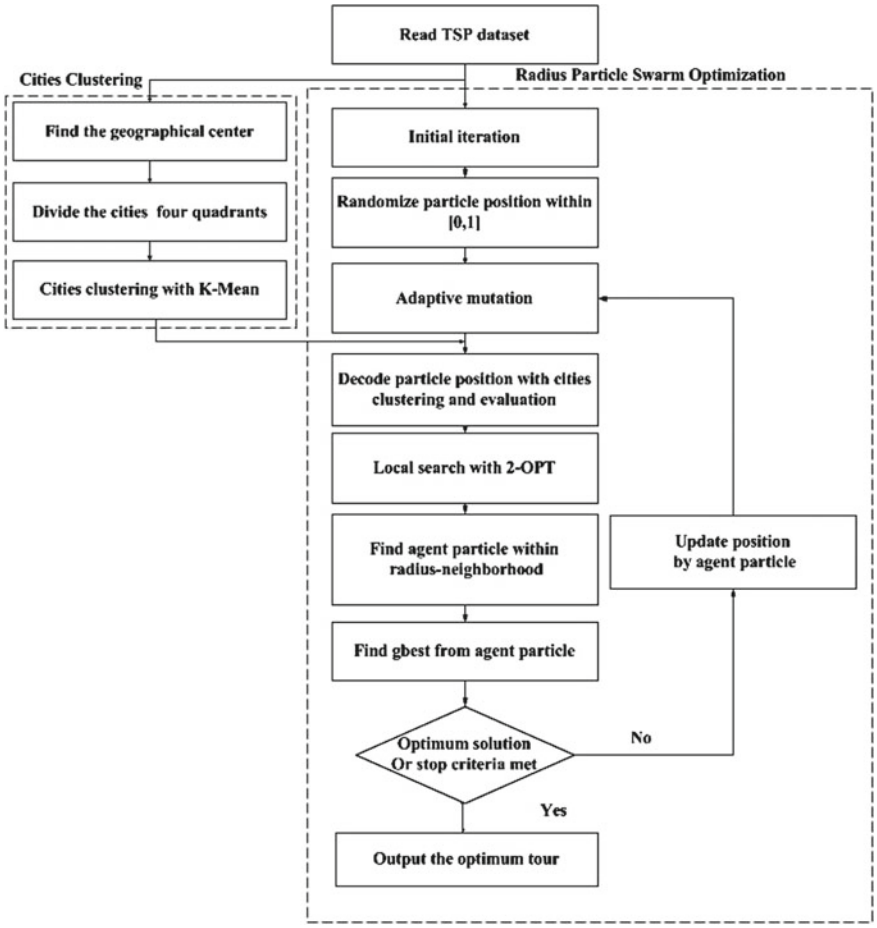


Fig. 5 The general framework of HRPSO for TSP

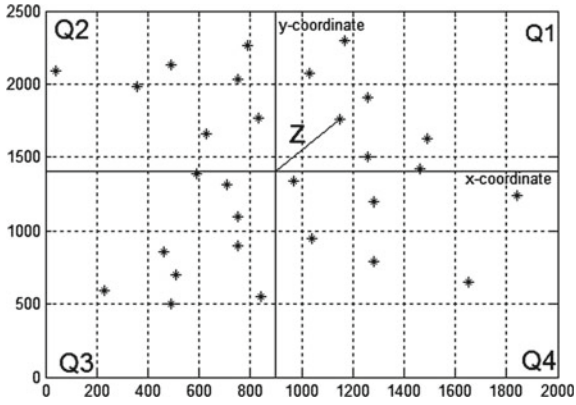


Fig. 6 The distribution of the cities (bays29) divided into four quadrants

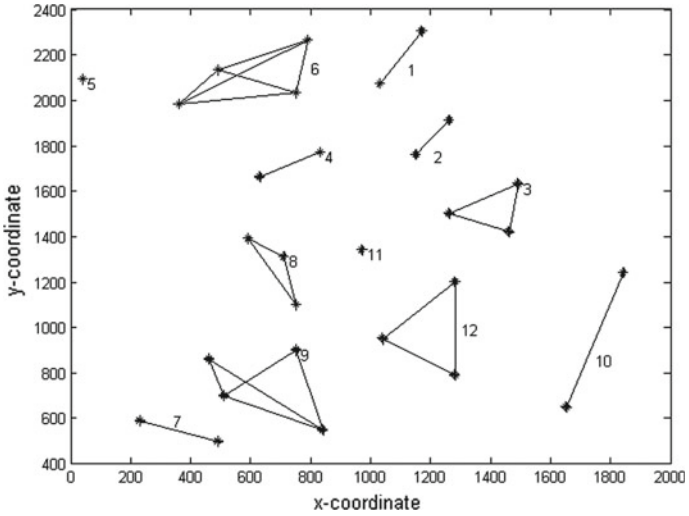


Fig. 7 The city clustering by k-means

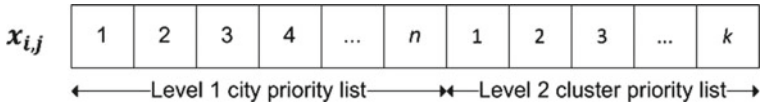


Fig. 8 The solution of TSP representation

created. Then, we randomly select k cities in the TSP dataset as initial centroids. After that, each city is assigned to the closest centroid. The location (x, y and z) of the centroid is determined as the nearest by the Euclidean distance in Eq. (7). This is generated for each cluster by re-computing the mean of the object set assigned to each cluster by Eq. (8). The iterative city clustering is repeated until clusters do not change. The city clustering is displayed in Fig. 7.

$$d = \sqrt{\left(c_x - \rho_{x_i}\right)^2 + \left(c_y - \rho_{y_i}\right)^2 + \left(c_z - \rho_{z_i}\right)^2} \tag{7}$$

$$\sigma(x, y, z) = \frac{1}{cn} \sum_{\sigma_i = 1; \forall \theta \in \sigma}^{cn} \theta_{\sigma_i} \tag{8}$$

3.2 Solution Representation

The solution is encoded using the priority values for each city. It consists of two levels, as shown in Fig. 8. There is the first level, which is associated with the city priority list and the second level, which is associated with the cluster priority list.

Therefore, the minimum value has the highest priority. The first level consists of n dimensions, with each dimension attributed to a city. The second level consists of k dimensions, with each dimension attributed to a cluster. The first level size is equal to the number of cities, while the second level size is equal to the number of clusters. The nearest merge cluster procedure will change the city start and city end of cluster A and cluster B and then merge them into a single cluster as shown in Fig. 9. It should be noted that in the TSP, the fitness of each particle is related to the expected length of the tour route. Since the problem is to find the minimum cost as shown in Fig. 10

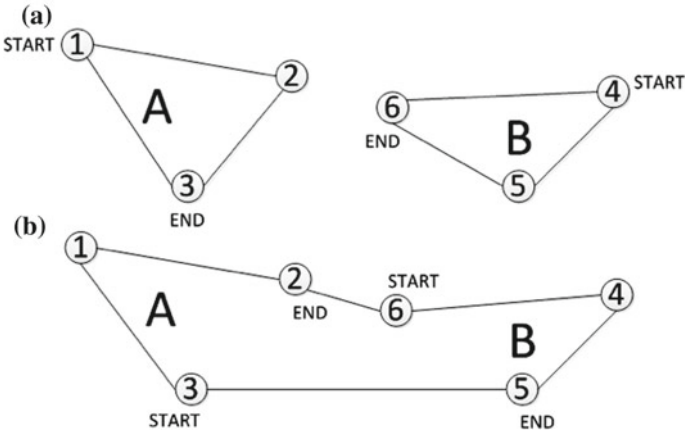


Fig. 9 The nearest merge cluster **a** sub-tours **b** the nearest merge tour

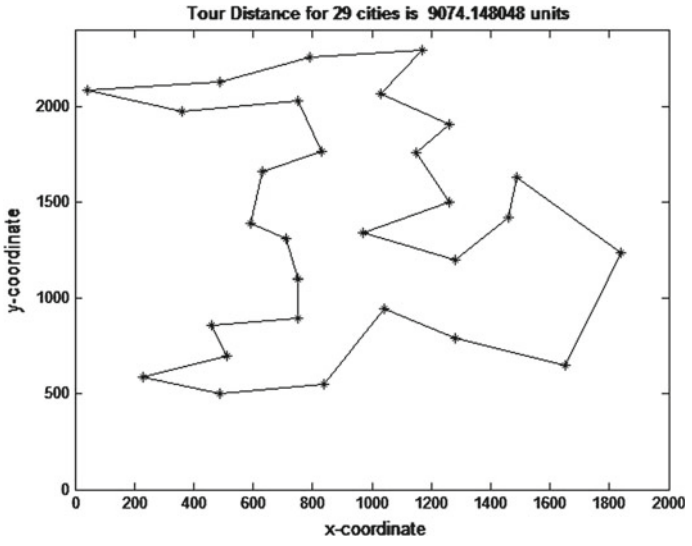


Fig. 10 The optimum tour for bayg29

Table 4 Comparison of average distance for Bayg29, Att48 and kroA100

Algorithm	bayg29	Att48	kroA100
HRPSO	**9074.18**	**10,648**	**21,341**
PSO	12,556.90	11,648	22,434
Route [11]	9336.00	–	–
CPSO [12]	–	–	21,689
MHPSO [13]	–	–	21,282
PSO [14]	–	10,867	22,071
GA [14]	–	10,828	21,851

(optimum tour of bayg29), when the global optimum is achieved or the procedure reaches the maximum number of iterations, the tour with minimum cost obtained by the k-means RPSO is returned as the output.

3.3 Experiments and Results

The proposed method is used to simulate the following three standard TSP datasets: Bayg29, Att48 and kroA100. The known optimal solutions are taken from the TSP Library at http://www.iwr.uni-heidelberg.de/groups/comopt/software/. The algorithm runs on the full datasets. We test our method against the traditional PSO algorithm. A total of 30 runs for each setting are conducted and average fitness of the best solution throughout the run is recorded. Therefore, the number of function evaluations is 60,000 (60 particles $\times$ 1000 iterations).

Several trials were performed to assess the effectiveness of the proposed method with different cluster number parameters that have value between 1 and 5. The proposed method provides the best solution for Bayg29 and Att48 with three clusters and on kA100 with five clusters. A summary of the optimum solution of the instances compare with other methods is given in Table 4.

4 Conclusion

The HRPSO solves the RCPSP by combining adaptive mutation and forward-backward improvement with RPSO. The adaptive mutation solves the impact on the maximum velocity or the minimum velocity in searching for the regional solution of the classical PSO so that the diversity among particles gradually decreases. The forward-backward improvement is employed to reassign the scheduling and reschedule the result acquired from the RPSO to improve convergence speed.

The k-means RPSO solves the Traveling Salesman Problem using the divide-and-conquer strategy. The method employs k-means algorithm to find the cities clustering and then solves a sequence of sub-cities in a given order and merge them by the Radius Particle Swarm Optimization.

The experiments show that the HRPSO provides the better performance to solve RCPSP and TSP than the traditional Particle Swarm Optimization and other existing methods. The performance of the HRPSO is tested against a number of instances from the PSPLIB and TSPLIB.

References

1. Boussaïd, I., Lepagnot, J., Siarry, P.: A survey on optimization metaheuristics. Inf. Sci. **237**, 82–117 (2013)
2. Parejo, J.A., Ruiz-Cortés, A., Lozano, S., Fernandez, P.: Metaheuristic optimization frameworks: a survey and benchmarking. Soft. Comput. **16**(3), 527–561 (2012)
3. Agarwal, A., Colak, S., Erenguc, S.: A neurogenetic approach for the resource-constrained project scheduling problem. Comput. Oper. Res. **38**, 44–50 (2011)
4. Hong Zhang, X.L., Li, H., Huang, F.: Particle swarm optimization-based schemes for resource-constrained project scheduling. Automation in Construction **14**, 393–404 (2005)
5. Bouzidi, S., Riffi, M.E., Bouzidi, A.: Comparative analysis of three metaheuristics for solving the travelling salesman problem. Trans. Mach. Learn. Artif. Intell. **5**(4) (2017)
6. Valls, V., Ballestin, F., Quintanilla, M.S.: Justification and RCPSP: a technique that pays. Eur. J. Oper. Res. **165**, 375–386 (2005)
7. Jia, Q., Seo, Y.: Solving resource-constrained project scheduling problems: conceptual validation of FLP formulation and efficient permutation-based ABC computation. Comput. Oper. Res. **40**, 2037–2050 (2013)
8. Ziarati, K., Akbari, R., Zeighami, V.: On the performance of bee algorithms for resource-constrained project scheduling problem. Appl. Soft Comput. **11**, 3721–3733 (2011)
9. Tseng, Lin-Yu., Chen, Shih-Chieh.: A hybrid metaheuristic for the resource-constrained project scheduling problem. Eur. J. Oper. Res. **175**, 707–721 (2006)
10. Valls, V., Ballestin, F., Quintanilla, M.S.: A hybrid genetic algorithm for the resource-constrained project scheduling problem. Eur. J. Oper. Res. **185**, 495–508 (2008)
11. He, X., Zhu, Y., Hu, H., Niu, B.: A new algorithm for TSP based on swarm intelligence. In: Proceedings of the 6th World Congress on Intelligent Control and Automation, pp. 3241–3244 (2006)
12. Yuan, Z., Yang, L., Wu, Y., Liao, L., Li, G.: Chaotic particle swarm optimization algorithm for traveling salesman problem. In: International Proceedings of the IEEE International Conference on Automation and Logistics, pp. 1121–1124 (2007)
13. Labed, S., Gherboudj, A., Chikhi, S.: A modified hybrid particle swarm optimization algorithm for solving the traveling salesmen problem. J. Theor. Appl. Inf. Technol., pp. 132–138 (2012)
14. Yan, X., Zhang, C., Luo, W., Li, W., Chen, W., Liu, H.: Solve traveling salesman problem using particle swarm optimization algorithm. Int. J. Comput. Sci., pp. 264–271 (2012)

Classification of Cell Nuclei Using CNN Features

Yuji Iwahori, Yuya Tsukada, Takashi Iwamoto, Kenji Funahashi, Jun Ueda and M. K. Bhuyan

Abstract This paper proposes a method to discriminate benign or malignant of cell image. Proposed method uses a CNN to extract features from the nuclei images detected from the original cell images and benign or malignant is automatically classified by two classes classification with SVM. This paper treats a cancer classification method using features obtained by using CNN from cell nuclei extracted from images of melanoma. The effectiveness of the proposed method has been confirmed by experiments from the viewpoint of cell classification of benign or malignant. It is shown that the difference between the number of cell nuclei diagnosed as normal and the number of cell nuclei diagnosed as cancer is clear in the benign and metastasis, while the difference in malignant is vague. Moreover, it is confirmed that N/C ratio of malignancy and metastasis was slightly higher than that of benign.

Keywords Cell Image · CNN · SVM · Classification · Benign · Malignant · Metastasis

Y. Iwahori (✉)
Department of Computer Science, Chubu University, 1200 Matsumoto-cho, Kasugai 487-8501, Japan
e-mail: iwahori@isc.chubu.ac.jp

Y. Tsukada · K. Funahashi
Department of Computer Science, Nagoya Institute of Technology, Gokiso-cho, Showa-ku, Nagoya 465-8555, Japan
e-mail: kenji@nitech.ac.jp

T. Iwamoto
College of Life and Health Sciences, Chubu University, 1200 Matsumoto-cho, Kasugai 487-8501, Japan
e-mail: iwamoto@isc.chubu.ac.jp

J. Ueda
Asahikawa Medical University, Asahikawa, Hokkaido 078-8510, Japan
e-mail: junueda@asahikawa-med.ac.jp

M. K. Bhuyan
Indian Institute of Technology Guwahati, Guwahati 781039, India
e-mail: mkb@iitg.ernet.in

R. Lee (ed.), *Computer and Information Science*, Studies in Computational Intelligence 849, https://doi.org/10.1007/978-3-030-25213-7_13

1 Introduction

In recent years, new clinical examination and imaging devices have been developed and improvement of diagnostic accuracy have led to the development of new therapeutic methods for various diseases in the field of medical pathological diagnosis. Although recent advances in such medical care are remarkable, cancer deaths are still increasing and the threat of cancer has been continued. Early detection of cancer is very essential to increase the chance of cure and to reduce mortality. Since cancer appears due to the occurrence of gene abnormalities according to the age, cytology directly observing cells plays a very important role in early detection of cancer in general. So the rapid and accurate pathological cytology based on data obtained from biopsy and exfoliation cytology is essential for the treatment of cancer. However, most of these diagnoses still depend on the experience and skill of the pathologist with time-consuming problem and quantification is difficult with poor objectivity. Therefore, there is a problem that may lead to an adverse result for patients that different diagnoses are made between hospitals. In addition to the problem that the number of cancer patients is increasing, the number of pathologists is only 2405 [1] in Japan at present and this is extremely low at 1 per 50,000 people. As a result, the diagnosis may not be sufficiently followed and there is some possibility that oversight or misdiagnosis may increase due to the increase of the load of pathologist.

From these situations, it is considered that the demand for an automatic diagnostic system for pathological diagnosis of cancer is increasing. Various tests and surveys [2, 4] have been conducted to develop an objective and quantitative automated diagnostic system for cancer due to the various forms of tumor cells in the body, but these trials have not been reached to the practical use yet. Reference [3] is proposed as a pathological diagnosis support system, where authors are improving the pathological diagnosis support system using Higher-order Local Auto Correlation (HLAC) features, This system is developed to make it easy for the pathologist to see the image for the diagnosis and the system itself cannot identify cancer and the load of the pathologist remains. The most important factor in constructing an automatic diagnosis system is the feature that determines whether it is cancer or not. Reference [4] provides a survey of feature quantities and feature quantities are examined for the automatic discrimination of colon cancer, but there is a problem of lack of practicability and accuracy such as manual extraction of images to extract some feature quantities with exclusion of few examples.

Although some machine learning approaches [5, 6] have been proposed, this paper proposes a cancer classification method using Convolutional Neural Network (CNN) features of cell nuclei extracted from images of melanoma, where a cancer diagnostic system is examined to reduce the burden on pathologists.

It is generally known that cell nuclei stained with various compounds are often used as markers for cytology in pathological diagnosis. Here the features of cell nuclei is focused and staining of cell nuclei by Hematoxylin-Eosin (HE) which is commonly introduced in pathological tissue specimens, is used to extract cell nuclei from tissue samples. This paper proposes a method to diagnose cancer by extracting

feature quantities from extracted cell nuclei using CNN and the paper evaluates the effectiveness for cancer diagnosis.

2 Background

This section describes details of staining of cell nuclei, machine learning and image features used in learning.

2.1 Hematoxylin and Eosin Staining

In the histopathological diagnosis using an optical microscope, it is necessary to color and stain originally colorless cells or tissues. Hematoxylin and eosin (HE) staining is used for the purpose of grasping the whole image of cell and tissue structure at the optical microscopic level. It is the most basic and important method of staining of pathological specimens. Although the staining results may differ slightly depending on the thickness of the section, the fixation conditions and the type of hematoxylin and eosin used, the nucleus is stained blue with hematoxylin, while cytoplasm/fibers and erythrocytes are colored pink with eosin as shown in Fig.1.

2.2 Support Vector Machine (SVM)

SVM is one of the methods of pattern recognition and inverse problem analysis, and basically it is a learning machine for solving classification problems of two class feature patterns. SVM is originally a linear classifier but can be extended nonlinearly by incorporating a kernel. In addition, SVM is extended to perform multi-class classification by applying two class classification multiple times.

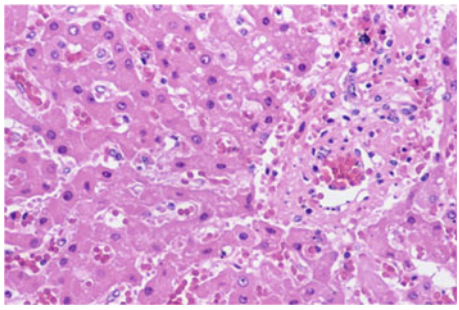


Fig. 1 HE stained pathological image

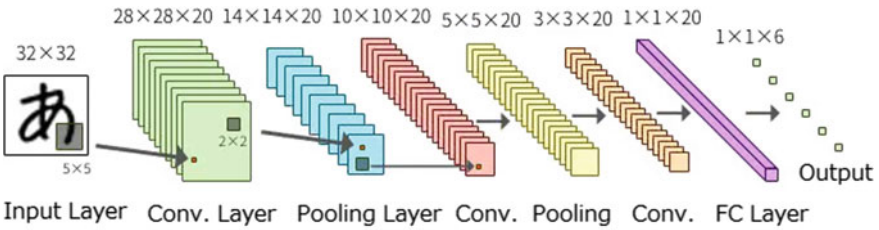


Fig. 2 Basic structure of CNN

2.3 Convolutional Neural Network

Convolutional Neural Network(CNN) has been used as Deep Learning algorithm and it has a high accuracy in the field of general image recognition. CNN consists of convolution, pooling and fully connected layers and CNN structure is given by repeating convolution process and pooling process and connecting to fully connected layer as shown in Fig.2. Most CNN models are developed and used for general object recognition problem and this paper uses the CNN for feature extraction to the problem of cell nuclei.

2.4 Features Obtained from AlexNet

CNN is used to classify the images of classification task as input to the defined layers originally. Learning is performed by updating the internal parameters of each layer but a large amount of data and a high specific computer are required for the learning. AlexNet [7] is a type of CNN model and this is a neural network that won at ILSVR 2012. This model is set with parameters as shown in Fig.3.

3 Classification of Cancer from Extracted Cell Nuclei

Here we describe a classifier used for cancer classification and a cancer classification program using the information of extracted cell nuclei.

3.1 Creating Classifier

To classify cancer from extracted cell nuclei, a classifier that determines whether it is cancer or not is necessary. Here, classifiers are explained for cancer classification and features that are judged to be cancer.

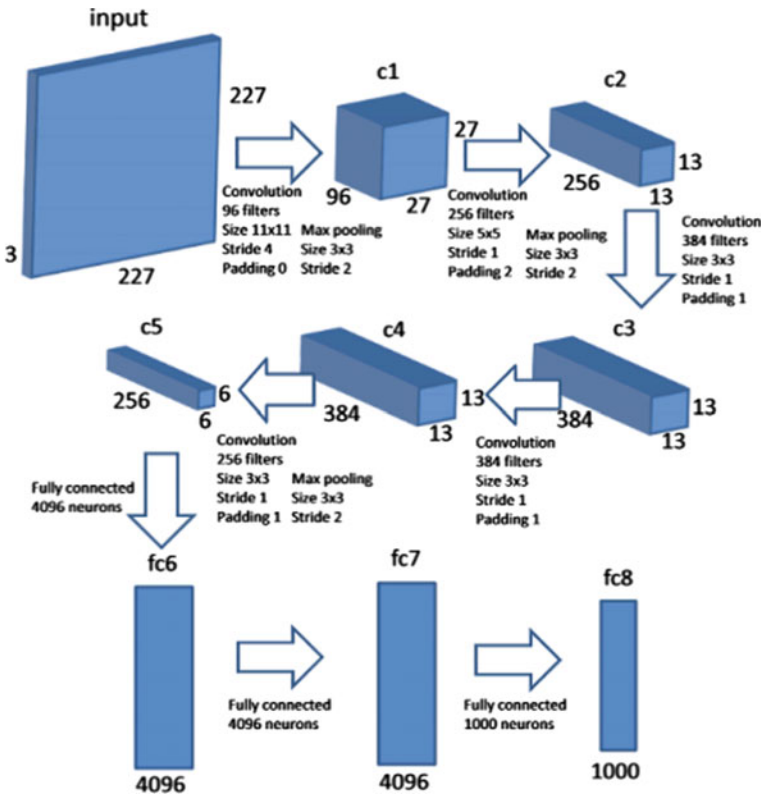
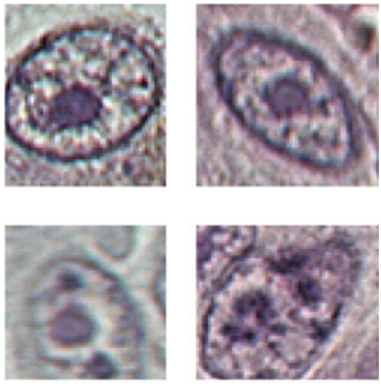


Fig. 3 Construction of AlexNet

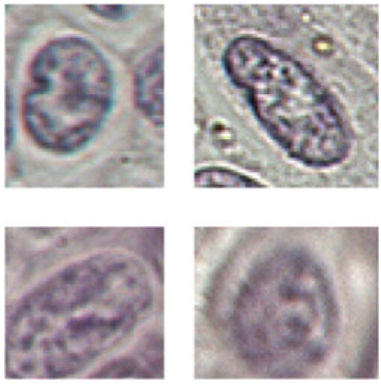
3.2 Creating Data Set

The degree of nuclear heterogeneity is one of the judgment factor for cancerous cell nuclei. Nuclear heterogeneity is generally associated with tumor grade (easiness to increase, spread), There is no specific indicator as to how large the degree of nuclear heterogeneity can be judged as cancerous. The nuclear heterogeneity of all nuclei does not necessarily increase and creating data set is more difficult than the case of cell nucleus extraction.

Here, the target of classification is melanoma, which is a type of skin cancer, Malignant cell nuclei with relatively large degree of nuclear heteromorphy from pathological image of melanoma. From the images of benign tumors such as lentigo, the nuclei with relatively small nuclear a typicality is labeled as normal, and data set is created. Fig. 4 is a part of the created data set, A total of 1600 images of cell nuclei consisting of 800 images labeled as malignant and 800 images of cell nucleus images labeled as normal were used as data sets.



(a) Large Cell of Nuclear Heterogeneity



(b) Small Cell of Nuclear Heterogeneity

Fig. 4 Data set used to classify cancer

3.3 Comparison Experiment of Features Used for Classification

It is necessary to extract features from the data set to classify cell nuclei. However, features suitable for expressing differences in nuclear heterogeneity are unknown at this stage. Therefore, an experiment is conducted using the created data set to confirm what kind of feature can correctly classify cancer or normal. Normal cell nuclei has large variations in brightness within the nuclei. It has been reported that the carcinogenic cell nucleus has a flat intensity variation in a whole.

We compared two of the histogrammed intensity [8] and the features automatically generated by CNN with relatively high accuracy in cell nucleus extraction. The data set used for evaluation was the same and 800 images labeled as malignant and 800 images classified as normal were used. As in the case of cell nucleus extraction, 70% of the data set is used for the learning, while 30% of the data set is used for

Table 1 Results of comparative experiments

Features	Accuracy
Intensity	0.723
CNN	0.891

the testing by 10-fold cross-validation in the evaluation. The evaluation value is the average value when cross-validation is performed 5 times for each feature value. Table 1 of the comparison experiment is shown below.

Table 1 suggests that CNN has the higher classification accuracy and CNN features are used for the following cancer classification.

4 Creating Cancer Classification Program

In this section, we describe a cancer classification program using CNN features. Classification of cancer is performed by the following procedure.

Step 1: Label the extracted cell nuclei.
Step 2: Make a rectangular region of cell nucleus.
Step 3: A rectangular cell nucleus region is used as an input image, and feature quantities are generated by CNN.
Step 4: Generated feature quantities are input to a classifier to classify whether it is malignant or normal by SVM.
Step 5: Repeat the processing from **Step 2** to **Step 4** until all cell nuclei have been classified.

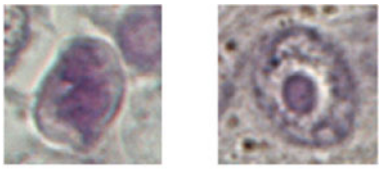
4.1 Rectangularization of Cell Nucleus Region

Extraction target must be an image as shown in Case 1 of Fig.5 to extract features by CNN. However, when the only the region of cell nucleus is extracted as shown in Case 2 of Fig.5, it is impossible to extract the feature by CNN as it is.

Therefore, the cell nucleus region should be rectangular so that the cell nucleus can be treated as an image. The following shows the procedure of rectangularization.

Step 1: The maximum value and minimum value of x coordinate of extracted cell nucleus and the maximum value and minimum value of y coordinate are derived respectively.
Step 2: The long side a, short side b and center coordinates of the rectangle are derived from the derived maximum and minimum values.
Step 3: In the case of $a \leq 90$, interpolate from the region around the input image so that the size of the rectangle is 90×90 with the center coordinates, as shown in Fig.6-(a).

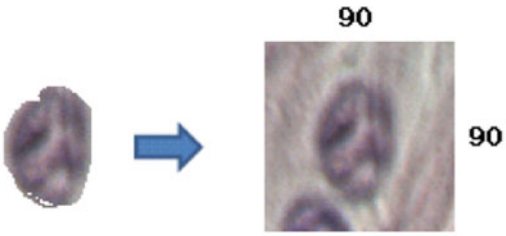
Fig. 5 Feature extraction by CNN



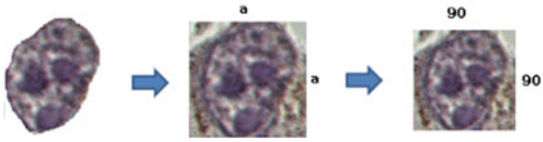
Case1: Feature Extraction is posssible by CNN

Case 2: Feature Extraction is imposssible by CNN

Fig. 6 Creating rectangle



(a) Case of a≤90



(b) Case of a>90

Step 3': If $a > 90$, create a rectangle with a × a and resize it to 90 × 90, as shown in Fig.6-(b).

5 Experiment of Cancer Classification

The effectiveness of the proposed method is confirmed through experiments of cancer classification using actual pathological images. Based on the various classification results, classification for each cell nucleus is considered.

5.1 Experimental Conditions

The pathological image obtained by HE staining of cells collected from human skin which a pathologist judged as a melanoma was used in the experiment. There are three types of judgment: "benign", "malignant" and "metastasis", "Benign" is diagnosed as benign tumor, "malignant" is diagnosed as malignant tumor. "Metastasis" is a type of malignancy among which it has already metastasized. The resolution of the HE stained image treated as an input image is 1920×1440.

Although there is no clear criteria such as benign or malignancy for "each cell nuclei" in the actual cytology, the classification of the test image was conducted as follows. If there were more cell nuclei judged to be benign among the whole extracted cell nuclei, the image was classified as normal, otherwise classified as cancer.

The judgment for the image was treated as correct when the labeled image as "normal" was classified to "benign" and labeled image as "cancer" was classified to "malignant or metastatic".

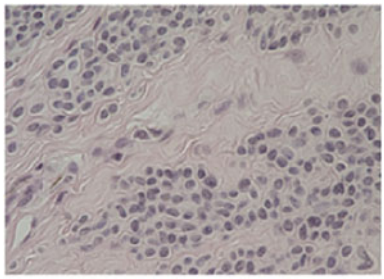
5.2 Classification Experiment

The pathological image is taken at high magnification with a magnifying microscope. Since several thousands of cell nuclei are scattered in the image, it takes time to classify cell nuclei one by one. Therefore, the pathology image and the section image obtained by dividing the pathology image into 16 equal parts are classified, We verified the possibility of classification for the case when part of the image was cut out.

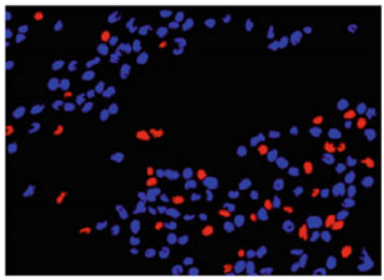
5.2.1 Classification Experiment in Each Section Image

The proposed method classified a total of 80 section images: 20 benign images, 40 malignant images and 20 metastatic images. The input images labeled as benign, malignant and metastas were used in the experiment and part of the classification results are shown in Figs.7, 8 and 9. In addition, the blue region in the classification result image is the cell nucleus diagnosed as normal, while the red region is the cell nucleus diagnosed as cancer. The number of cell nuclei classified as normal in each image, the number of cell nuclei classified as malignant in each image, the classification results, labels, and the overall accuracy of correct classifications are shown in Table 2.

Table 2 suggests that the overall accuracy of the classification gives 0.82 and the effectiveness of the proposed method was confirmed. In particular, there are many cell nuclei labeled as "Benign" and diagnosed as "normal", and many as cell nuclei labeled as "Metastasis" and diagnosed as "cancer" and the difference between these two kinds is clearly confirmed. However, there were some misclassifications for the

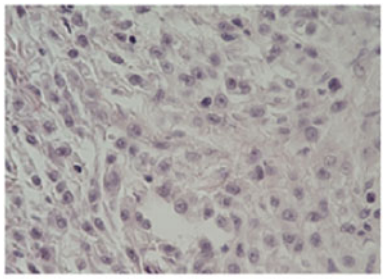


(a) Input Image

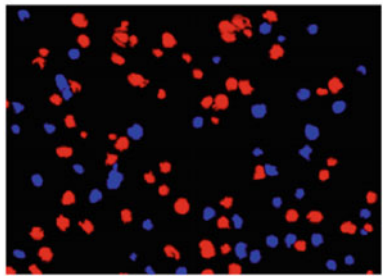


(b) Classification Result

Fig. 7 Images of benign tumors

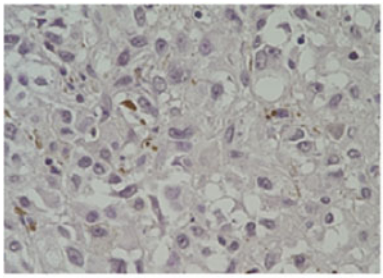


(a) Input Image

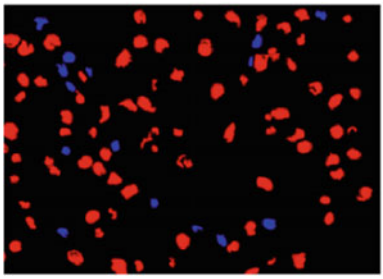


(b) Classification Result

Fig. 8 Image of malignant tumor



(a) Input Image



(b) Classification Result

Fig. 9 Image of metastatic cancer

Table 2 Results of classification experiments on section images

Label	Benign classifiers	Malignant classifires	Accuracy	Whole accuracy
Benign	20	0	1.0	0.82
Malignant	15	25	0.63	
Metastasis	0	20	1.0	

malignancies since there are many cell nuclei that have been diagnosed as “normal”. Results were based on the situation that number of “normal” labeled cell nuclei was larger and those images were misclassified or that the difference did not appear so much in comparison with the two cases (benign and metastasis) and results for the malignant were not so effective as a result.

5.2.2 Classification Experiment for Whole Pathological Image

Proposed method classified the pathological images for a total of six images: two benign images, two malignant images, and two metastatic images. Some of the benign, malignant, and metastatic pathological images used in the experiment are

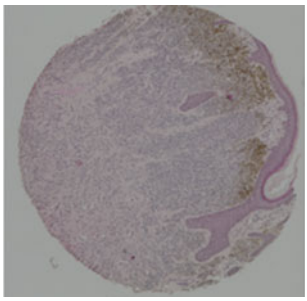


Fig. 10 Whole of benign tumor pathology image

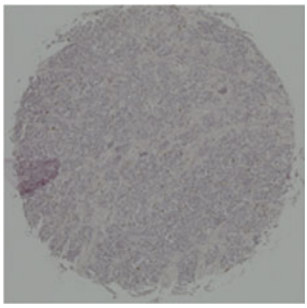


Fig. 11 Whole of malignancy pathology image

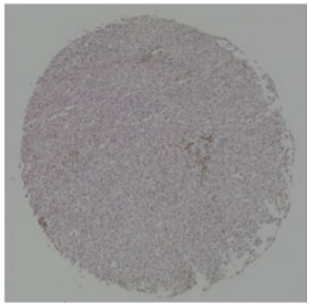


Fig. 12 Whole of metastatic cancer pathological image

shown in Figs.10, 11 and 12. Since the accuracy rate is not helpful because the total number of images used is small, the number of cell nuclei classified as normal, the number of cell nuclei classified as cancer and classification results, N/C ratio which is defined as the area ratio of nuclei to the cytoplasm was also obtained as a criteria for evaluation factor of the effectiveness. The value can be obtained by dividing the area of nuclei by cytoplasm. N/C ratio tends to become large because the area of the nucleus increases in tumor cells, especially in cancer. Therefore, N/C ratio is one of the parameters used by pathologists as an evaluation index for cancer diagnosis. The results are shown in Table 3.

Table 3 suggests that the difference between the number of cell nuclei diagnosed as normal and the number of cell nuclei diagnosed as cancer is clear in the benign and

Table 3 Results of classification experiments on pathological images

Label	Number	Benign nuclei	Malignant nuclei	N/C ratio	Classification result
Benign	1	1780	581	0.23	Normal
	2	1869	430	0.22	Normal
Malignant	1	1089	866	0.28	Normal
	2	813	874	0.28	Cancer
Metastasis	1	499	1814	0.31	Cancer
	2	421	1905	0.30	Cancer

metastasis, while the difference in malignant is vague. Moreover, it is confirmed that N/C ratios of malignancy and metastasis were slightly higher than that of benign.

5.3 Discussion on Cancer Classification by Each Cell Nucleus

The actual pathological diagnosis is done by observing the N/C ratio and the entire image of cell nuclei should be considered with competitive learning [9, 10] based on the global observations, and classification judgment is not performed for each cell nuclei like proposed method, where proposed method tried to judge each cell nuclei based on the the nuclear heterogeneity of cell nuclei. From the classification accuracy, the effectiveness of the proposed method was confirmed. Also, based on the results of two types of experiments, the proposed method using CNN feature and SVM can classify correctly with respect to two of "benign" and "metastasis", while "malignant" is not so accurate including the case that normal and malignant cell nuclei are almost evenly distributed. Although the specific stage of the data set used for this experiment is unknown, we can see that the case of "Metastasis" is higher from the N/C ratio of the experimental results. The nucleus of the image labeled "Metastasis" is considered to have a higher stage than "Malignant".

6 Conclusion

This paper proposed a cancer classification method using features obtained by using CNN from cell nuclei extracted for melanoma. The effectiveness of the proposed method has been confirmed with the experimental results from the viewpoint of cell status and the stage of cancer.

Development of a method to separate overlapping parts from extracted cell nuclei is remained as a futher problem. Although the proposed method used CNN and SVM but classification using Fine-tuning by CNN and classification taking into account the stage of cancer are also remained.

Acknowledgements This research is supported by JSPS Grant-in-Aid for Scientific Research (C) (17K00252) and Chubu University Grant.

References

1. Japanese Pathological Association: List of accredited pathologists as of October 1, 2017 (in Japanese)
2. Taylor, C.R., Levenson, R.M.: Quantification of immunohistochemistry–issues concerning methods, utility and semiquantitative assessment II. Histopathology **49**, 411–424 (2006)
3. Kurihara, T., Nori, H. *et al.*: Abnormality detection method from pathological images using higher order local autocorrelation features (in Japanese). Research Report of Mathematical Modeling and Problem Solving (MPS) 2010-MPS-81 (32), pp. 1–6 (2010)
4. Harai, Yuriko, Tanaka, Toshiyuki: Feature extraction for automatic determination of colorectal cancer grade (in Japanese). Proc. Jpn. Joint Autom. Control Conf. **57**, 1105–1107 (2014)
5. Koza, J.R.: Genetic programming, on the programming of computers by means of natural selection. MIT Press (1992)
6. Metropolis, Nicholas, Rosenbluth, Arianna W., Rosenbluth, Marshall N., Teller, Augusta H., Teller, Edward: Equation of state calculation by fast computing machines. J. Chem. Phys. **21**, 1087 (1953)
7. 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)
8. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition(CVPR), pp. 886–893 (2005)
9. Amoyel, Marc, Bach, Erika A.: Cell competition: how to eliminate your neighbours. Development 2014 **2014**(141), 988–1000 (2014)
10. Hogan, C., Kajita, M., Lawrenson, K., Fujita, Y.: Interactions between normal and transformed epithelial cells: their contributions to tumourigenesis. Int. J. Biochem. Cell. Biol. **43**(4), 496–503 (2011)

An Exploratory Study on the Benefit Sharing of Broadcasting Content Industry

Seung-Hyeok Baek, Min-Gwon Park, Jin-Hua Zhang and Jun-Ho Lee

Abstract Recently, it is becoming increasingly important to establish and utilize desirable inter-organizational relationship in the broadcasting contents industry. This study examined benefit sharing between organizations as a way to strengthen competitiveness in terms of co-evolution through inter-organizational relations in the Korea's broadcasting contents industry. Especially, it attempted an exploratory approach to identify the current status of benefit sharing (especially in terms of revenue sharing) between OTT (Over The Top), a new distributor, and Content Providers (CPs) and ways of benefit sharing, as well as the relationship between traditional TV broadcasters and CPs. The study also dealt with the issue of benefit sharing for co-evolution through discussion of the typology of benefit sharing, including OTT, which represents the new broadcasting contents market structure. In addition, the study also presented its limitations and proposed a direction for future research.

Keywords Broadcasting content · Co-evolution · Benefit sharing

1 Introduction

Benefit sharing is an essential and important issue in inter-organizational transactions. As rapid and discontinuous changes become commonplace in recent years, it is virtually impossible for a single organization to gain and sustain a competitive advantage on its own. Under these circumstances, network organization, which is actively discussed these days, may be considered an alternative means of utilizing inter-organizational relationships to gain a competitive edge. However, while applying network organization to every organization is somewhat impracticable, benefit

S.-H. Baek
Policy Division Future Policy Team, Korea Creative Content Agency, Naju-si, Korea

M.-G. Park
Department of Business Administration, Hongik University, Mapo-gu, Seoul, Korea

J.-H. Zhang · J.-H. Lee (✉)
Department of Business Administration, Hoseo University, Cheonan, Chungnam, Korea
e-mail: jhpaul2011@naver.com

R. Lee (ed.), *Computer and Information Science*, Studies in Computational Intelligence 849, https://doi.org/10.1007/978-3-030-25213-7_14

sharing can be utilized as a much more generalized means in relationships between organizations or between any organization and its employees.

Regardless of industries, benefit sharing is significant in most inter-organizational transactions, and even more so when it comes to the broadcasting content industry, especially the Korean broadcasting content industry. Presently, the nation's broadcasting content industry is at a crossroads as to whether it takes another leap or faces a decline in the functional relationship between various external and internal factors. Creative players in the country's broadcasting content industry have contributed to the mass production, distribution and globalization of attractive contents called 'Hallyu, or the Korean wave,' attracting great interest in and expectations for the industry.

While drastic and discontinuous alterations in commodities, services and technology have gradually accumulated in other industries, the broadcasting content industry inherently undergoes frequent and dramatic changes. Thus, there is a great deal of concern at home and abroad as to whether benefits that match such interest and expectations will be steadily guaranteed in the future as well. Against this backdrop, it is very crucial to share benefits, instead of enhancing competitiveness of each individual organization, in order to strengthen inter-organizational relationships as a source of competitiveness to achieve stability, efficiency, and innovation.

As to the Korean broadcasting content industry, the need for traditional benefit sharing has arisen in social terms rather than in economic terms in the relationship between contents providers (CPs) in poor production environments and affluent terrestrial broadcasters monopolizing distribution. Responding to such need, various government authorities have intervened and made efforts as well. In other words, it can be said that terrestrial broadcasters had no driver for sharing benefits with CPs, except for social values. In a monopolistic distribution structure in which CPs could not produce profits through any other routes, their poor conditions made things worse.

However, major changes have occurred in the nation's broadcasting content industry in recent years, presenting a task to utilize benefit sharing as a driving force for sustainable development in economic terms, particularly in terms of co-evolution. More specifically, two significant changes have occurred. One is channel diversification. The number of TV channels, including CATV and comprehensive programming channels, has relatively continuously risen. Also, distribution networks, such as OTT and one person channels which are spreading through Netflix and YouTube, have increased. This phenomenon has brought about various forms of transactions beyond the traditional relationship between CPs and terrestrial broadcasters, raising the issue of benefit sharing. In particular, Netflix's approach to benefit sharing has drawn a lot of interest and attention and affected the industry in recent years.

As the other main change, the boundary between distribution and production has become blurred. In other words, organizations (CJ ENM, Netflix, Kakao Story, etc.) previously engaged in distribution have advanced into the production area. This has raised the issue of benefit sharing between production and distribution in a

loosely-coupled context, which is different from benefit sharing on the assumption of transactions between the existing CPs and master contents providers (MCPs) or from traditional vertical integration.

Regarding these two major changes in the Korean broadcasting content industry, this study will focus on the former because it is rather premature to discuss benefit sharing in relation to the latter. Specifically, we will have exploratory discussions about the direction for sharing benefits in which TV, a traditional channel among diversified channels, and OTT, among new channels, accomplish co-evolution through their relationships with CPs in the domestic broadcasting content industry, and then we will make suggestions.

Particularly with regard to benefit sharing between CPs and MCPs, we will seek many best ways according to circumstances in terms of contingencies, rather than seeking one best way from a universal point of view. In this way, we will attempt exploratory typology for types of benefit sharing between CPs and MCPs, hopefully contributing to facilitating future studies and discussions. However, types of benefit sharing in the broadcasting content market are not diverse. As pointed out earlier, the relationship between terrestrial broadcasters without the driver for benefit sharing and CPs failing to gain an edge in negotiating power has lasted for long. As a result, compared with the manufacturing market which has considered various ways of sharing benefits through cash, stocks, etc., types of benefit sharing in the broadcasting content industry are monotonous. Benefits have therefore been shared mainly through revenue sharing in the industry. Hence, we will discuss benefit sharing in terms of revenue sharing.

2 Benefit Sharing and Revenue Sharing

2.1 Benefit Sharing

It can be said that the notion of benefit sharing began in the manufacturing industry which is the basis of industry. With the concept of outsourcing established and with entities among industries diversified, mutual benefits have begun to be considered significant. With the enactment of the ‘Act on the Promotion of Collaborative Cooperation between Large Enterprises and Small-Medium Enterprises,’ benefit sharing between the outsourcer and the oursourcee has become institutionalized (March 3, 2006). Under Article 8 of the same Act, benefit sharing is a “contract model under which outsourcers support outsourcees to attain their common and agreed goals, including the cost-cutting, etc. and both the outsourcers and outsourcees share the results.” In August, 2007, the concept of benefit sharing between enterprises and their employees was introduced in the ‘Special Act on Support for Human Resources of Small and Medium Enterprises’ so as to facilitate benefit sharing between them. In other words, the institutionalization of benefit sharing and between enterprises and their employees has laid the foundation for their mutual growth.

As mentioned above, however, the introduction of benefit sharing in the broadcasting content industry is too slow, which is found in between enterprises and between enterprises and their employees. This phenomenon mainly stems from the difference in bargaining power of broadcasters and CPs. The completeness and quality of contents do not always guarantee success. Even high-quality contents become successful through distribution. This unique nature of contents has helped maintain the vested interests of broadcasting companies. According to Broadcasting Industry White Paper 2017 [1], the total number of broadcasting business operators and CPs stands at 957 in 2016, up 0.3% from a year earlier. In this data, the broadcasting business operators mean business operators in five industries, namely, terrestrial broadcasting, cable broadcasting, satellite broadcasting, DMB, and IPTV, and no big difference appears in the number of companies in these industries. In other words, alterations in the number of companies occur largely due to an increase or decrease in CPs. As to CPs, small enterprises with fewer than 10 employees account for 50.5% of them. This structure inevitably brings about a difference in negotiating power of broadcasters and CPs.

2.2 Revenue Sharing

One type of benefit sharing is to share revenues (profits). As for the manufacturing industry, benefit sharing is largely divided into four types depending on how companies share profits or benefits: cash, stocks, funds, and corporate partnership [2].

As for the content industry, however, only a small proportion of companies among approximately 105,000 enterprises are listed companies. They lack shared assets and thus, they mostly share revenues to share benefits.

The concept of revenue sharing is not limited to the content industry. Rather, the notion has been regarded as an important issue that forms the basis of industry since economic activities occurred. Also, sharing became a standard for exchange even before the introduction of the economy. Various forms of sharing have existed and economic entities have endeavored to establish a reasonable standard for sharing revenues. If at least two transaction objects are involved, sharing must occur. Success or failure depends on whether the standard and sharing are reasonable, and sustainability of economic activity is maintained only when they are reasonable.

Nevertheless, the standard for revenue sharing has caused controversy constantly. That is because there is a difference in bargaining power of entities participating in economic activities under the same goal. Negotiating power alters a standard for revenue sharing, and it is difficult to create a consensus that satisfies all. In many cases, one party is forced to sacrifice itself. If such sacrifice continues, unfair practices begin to be used in the market. An entity that gains an advantage in negotiations quickly accumulates wealth and expands its market. In contrast, an entity in the opposite position often fails to achieve satisfactory results relative to economic resources invested, such as labor (planning, creation, production, etc.) and financial resources.

This often causes unfair practices that have recently been pointed out as harm to the industry. As market competition spreads globally beyond the domestic market, competition with business operators overseas is becoming a problem as well.

3 Types of Revenue Sharing

3.1 Changes in the Broadcasting Content Distribution Structure

These days, viewers have a wide choice of broadcasting contents. This has been possible because diverse routes have come out for viewers to enjoy broadcasting contents with the improvement of CATV, growth of IPTV, and advent of new distribution platforms, including Over-The-Top (OTT). Also, with the development of digital technology, technical perfection of similar broadcasting services led by mobile communication service providers has increased. As a result, the strict boundaries of broadcasting and communication have been blurred and the range of broadcasting content market is expanding to online and mobile.

In the past, broadcasting contents were produced by terrestrial broadcasters and broadcasting companies represented by program providers (PPs), organized by terrestrial broadcasters and PPs, and then transmitted through CATV, satellite broadcasting, and IPTV [3]. In 2016, however, the frequency of using broadcasting contents over mobile has surpassed that of TV. According to Broadcasting Industry White Paper 2017, the percentage of people using TV more than five days a week stood at 75.4% in 2016, while the percentage of people using smartphones more than five days a week stood at 81.0%. Also, TV viewing time has gradually decreased to less than three hours, or two hours and 57 min on average, since 2015.

These changes in viewing behavior and the broadcasting market have altered the competitive terrain of the broadcasting content market. In January, 2016, Netflix, a global OTT operator, began to provide service in Korea. Netflix's market impact is still small, but concerns about its potential market impacts based on its capital and distribution power as a global company are very significant. Mobile services which have previously started with additional services of broadcasting companies are increasing. Existing broadcasters, such as terrestrial broadcasters, PPs, and CATV operators, are also expanding their services into the mobile realm. This phenomenon is causing excessive competition not only over broadcast contents but also between video services. As they expand business to online distribution business, competition in the overall broadcasting distribution market is shifting to the mobile area.

3.2 Issue of Traditional Revenue Sharing

So far, many studies on broadcasting contents have focused on discussions on rectifying unfair trade (revenue sharing) in the broadcasting market and on promoting sharing following the emergence of new media and changes in viewing behavior [4]. Nevertheless, no specific methodology to do so has been come up with. As stated earlier, because the market is based on autonomous competition and priority is given to private contracts when it comes to contracts, the government's policy has served simply as a recommendation and thus contracts have hardly fulfilled their functions. However, it is crucial to set a standard for revenue sharing because sharing reasonable revenues works as a great inducement to plan new contents and to expand business models.

Baek et al. [5] present a standard for sharing based on the shares of contribution as a way to activate revenue sharing. They think that it is a reasonable standard to identify key participating entities from the stage of production to distribution of broadcasting contents and to share revenues by measuring each participating entity's share of contribution to production and distribution. In the process of producing broadcasting contents, various entities participate, such as a service provider (SP) offering the final result to the viewers, a CP, a venture capital firm (VC), and an MCP. Also, both the shares of contribution and bargaining power of participating entities vary depending on their roles. Baek et al. [5] propose six types of contracts: ① CP-MCP agency contract; ② CP-MCP MG + RS contract; ③ CP-VC-MCP agency contract; ④ CP-MCP direct investment + (MG+) RS contract; ⑤ CP-VC-MCP agency contract; and ⑥ CP-SP direct contract.[1]

This revenue sharing based on the shares of contribution according to contract types is similar to that of the 'Production Committee' in Japan. In the country, multiple investors, including publishers, game companies, broadcasters, CPs (including labor investment), and record companies, often invest in producing broadcasting contents. After the finished content is exposed to the media, revenues are shared based on participating entities' shares of contribution to the production, and they own relevant copyrights to earn revenues from subsequent distribution channels [5].

However, the revenue sharing based on shares of contribution has not been established in the market. Both the final window (media) exposed to viewers and financial contribution to production costs are still highly appreciated. Therefore, when it comes to traditional revenue sharing, the bargaining power of an entity that invests in the media and production costs is recognized as significant.

Recently, signs of change have begun to appear in traditional revenue sharing. This is a result of the narrowing of the existing broadcasters' position as mobile has become the mainstream of broadcasting content distribution channels. In addition, significant changes are occurring in competition between terrestrial broadcasters and CATV. A free-to-air broadcast channel, CATV is free of government regulation, compared to terrestrial broadcasters. Therefore, CATV makes plans and various attempts more freely, getting a good response from viewers. Previously, viewing rates have been

[1] MG: Minimum Guarantee/RS: Revenue Sharing.

used as absolute data for judging broadcasting contents and attracting advertisements, but now the center of gravity has shifted to 'topicality.' Terrestrial broadcasters have also abandoned their habits of focusing too much on viewing rates and instead, devoted their efforts to planning and production that can bring about topicality.[2]

3.3 Issue of Revenue Sharing Focused on OTT Services

3.3.1 Current Status of OTT Services

Domestic OTT operators provided VOD-type contents online at the initial stage. 'Pandora TV,' which first appeared in October 2004, started as a video-sharing site that uploaded VOD service as well as video contents provided by one-person creators in user created contents (UCC) forms. In 2007, Pandora TV also introduced the world's first business model to distribute revenue to UCC uploaders before YouTube. After that, various services, such as 'Afreeca TV' (2006), 'Gom TV' (2006), and 'Daum tvPot' (2007), started to appear. With OTT services based on broadcasting contents (tving) launched in 2010, various types of service providers appeared, including those based on one-person creators and those based on broadcasting companies.

As for domestic OTT services, it is estimated that the top 13 services account for at least 97% of the total sales. Among them, 'YouTube' is most frequently used, followed by 'Facebook' and 'Naver TVcast' (current 'Naver TV'). As to usage rates, the usage rate of 'YouTube' is 93.2%, 'Facebook', 29.6%, and 'Naver TVcast', 18.3%. Regarding paid services, service providers with a relatively high usage rate include 'oksusu' (SKB), 'olleh tv mobile' (KT) and 'video portal' (LGU +) provided by the big three telecom companies; 'pooq' offered by the contents alliance of terrestrial broadcasters; and 'tving' provided by CJ ENM. Also, the percentage of users who enjoy the free services 'YouTube' (86.1%) and 'Naver TVcast' (57.7%) is high. In the case of paid services, the usage rates of 'Netflix' (74%) and 'pooq' (63.9%) are high. Among OTT service users, paid users take up 85.2%, which is higher than free users (60.9%) [6] (Table 1).

These OTT services are highly regarded as those that can replace existing content platforms. They can stably offer types of contents preferred by viewers and are also suitable for a changing trend in which media usage behavior is moving to the mobile realm. In addition, OTT operators, like Netflix, are establishing vertical business models on production and distribution to produce their own original contents, expanding their influence to the production and distribution markets and even to the advertising market [7].

[2]Terrestrial broadcasters are paying more attention to topicality than viewing rates, increasing the production of such programs as School Attack 2018, Dancing High, Visiting Teacher, After School Hip-Hop, and High School Rapper 2.

Table 1 Current status of OTT services

Classification	Service name	Service provider	Service year	Major service
Group1	Tving	CJ ENM	2010	Real time broadcasting for CJ's CATV programs, as well as VOD service (back view, including movies)
	Pooq	3 terrestrial broadcasters, EBS	2012	Realtime broadcasting for programs provided by terrestrial broadcasters and comprehensive programming channels, as well as VOD service (back view)
Group 2	Olleh tv mobile	KT	2011	Realtime broadcasting for programs provided by comprehensive programming channels + paid channels + overseas channels, VOD movies (back view)
	U+ mobile TV	LGU+	2014	Films, major TV programs of terrestrial broadcasters
	Oksusu	SKB	2016	Realtime broadcasting for programs provided by comprehensive programming channels + paid channels + overseas channels, VOD movies (back view) *This service was previously provided through 'Btv mobile' but integrated with hoppin service in 2011, and then replaced with the current service in 2016.
Group 3	Everyon TV	HCN	2011	PP for CATV and comprehensive programming channels, as well as independent channels
Group 4	Pandora TV	Pandora	2004 (first)	UCC, VOD service for existing broadcasting contents, etc.
	Gom TV	Gom and Company	2006	
	YouTube	Google	2008	
	Naver TV	Naver	2013	Broadcasting and web-only contents, UCC, existing broadcasting contents *In January 2017, the service name was changed from 'Naver TVcast' to 'Naver TV.'
	Kakao TV	Daum	2015	

(continued)

Table 1 (continued)

Classification	Service name	Service provider	Service year	Major service
Group 5	Afreeca TV	Afreeca TV	2006	
	Netflix	Netflix	2016	VOD movies, etc. (back view, monthly subscription), original contents
	Watcha play	Frograms	2012	Movie recommendation service ('watcha' recommendation algorithm), VOD movie and soap opera streaming service

* Each OTT operator's business models, contents provided, and service methods are not fixed but rather constantly changing and evolving, and no strict academic standards exist yet. For such reason, this typology is not based on absolute standards
* *Source* The Korea Communications Commission (2017), p. 274, reorganized

Therefore, it is very crucial to deal with the issue of revenue sharing (benefit sharing), focusing on OTT services. That is because a new standard for sharing can be established when a new business model appears.

3.3.2 Broadcasting Content Type

So far, the success of introduction of new media has been heavily dependent upon broadcasting contents. New technologies provide viewers with new experiences, but they will have no value as a medium if they cannot create a high-quality lineup of contents. Utilizing the contents of terrestrial broadcasters, in particular, has been suggested as a good way to connect with the audiences. Even when the first cable television service was launched in 1985, it pitched audiences on its multi-channel services. Yet, there was very little contents that the viewers enjoyed, except the content of terrestrial broadcasters. Over 100 other channels offering cheap foreign contents or out-of-date movies quickly fell out of favor with the viewers. DMB, which had been grabbing attention as mobile broadcasting, struggled to meet the demand with contents provided by terrestrial broadcasters; IPTV also relied on re-transmitting terrestrial broadcasters' contents to promote its service at an early stage. In other words, a basic business model is to boost service usage share by maintaining viewers' media experience with the supply of terrestrial broadcasters' contents and also by providing them with new experiences yet to be had with broadcasts via VODs and games.

Domestic OTT services can be broadly classified into three types: One creator-based OTT service, such as 'Afreeca TV' and 'Pandora TV'; OTT service based on broadcasting contents, such as 'pooq' and 'tving'; and an original content-based OTT service, such as 'Netflix.'

From a revenue sharing (benefit sharing) perspective, broadcasting content-based OTT and original content-based OTT are worthy of attention as they are based on the participation of broadcasters and CPs, two key players in the market for producing and distributing broadcasting contents.

The OTT platform's business model can be broadly divided into five categories. First, it is a Transactional VOD (TVOD) model, a service by which the consumer can purchase contents separately. It is the business model of Apple 'iTunes.' Second, a subscription VOD (SVOD), by which the subscribers pay a certain amount each month or year. Usually, the customer can rent content, which expires in 48 h, or obtain lifetime streaming rentals that can be accessed online forever with a single purchase. Subscribers can use the service, regardless of the number of times it is used, once they have paid a certain amount. Although many vendors have adopted the SVOD service model, the domestic vendors' sales are still negligible. The business model of 'Netflix' and 'Watcha' which adopts 100% SVOD model is the most glaring example of success. Third, it is a Advertising VOD (AVOD), a VOD service based on advertisement revenue. Although such business model is not widely adopted in Korea except on some big-name portal sites such as 'Naver' and 'Daum,' Pandora TV in Korea and giant OTT operators abroad, such as 'YouTube' and 'Hulu,' have adopted this model. Fourth one is a Free VOD (FVOD), which is a service available free of charge to anyone who signed up for the service. The FVOD service type is mainly used for launching new platforms or for attracting subscribers in a short period of time. Fifth one is a hybrid VOD service method operating based on the adoption of two or more business models at the same time from among the above-mentioned business models. A prime example is "YouTube." It is also a business model of the domestic telecom operators' OTT. These business models are emerging across video contents, including movies, as well as broadcasting contents (Table 2).

The OTT service offered by three major domestic telecom operators is a prime example of the OTT service based on broadcasting contents. There are a plurality of contract types depending on the contract conditions, the service range, the period, etc. for the contents to be supplied and purchased, resulting in different revenues. Usual contract types include ① flat sales (flat contracts),[3] ② contracts (RS contracts) based on an amount of sales, or ③ MG + RS contracts, which distribute revenue at a fixed rate based on sales in return for the minimum guarantee (MG).[4]

There is no agreed-upon industry standard for the distribution of revenue generated by these contract types, nor is there consensus or standardization among stakeholders. Depending on which one has bargaining power, the distribution rate is changed and a corresponding difference in revenue arises. Therefore, it is difficult to universally define a form of a clear market contract. Yet, it is easy to predict that even if the OTT service grows, that won't immediately make situations any better for CPs. As noted

[3] While it is not known how an RS contract involving broadcasting contents is usually concluded, in the case of movie contents, which is the same video content as broadcasting content, a flat contract is concluded based on estimation of sales expected to be generated within the contract period, reflecting anticipated sales amount or some 1.5 times that [8].

[4] For information on detailed contract types, please refer to the section entitled '2) Issue of Traditional Revenue Sharing.'

Table 2 Types of Business Models for Video Contents Services

Classification	Features	Key services
TVOD (Transactional VOD)	– A service that allows for separate purchase of content – (Usually) offering limited 48 h access to content/lifetime streaming rentals	iTunes
SVOD (Subscription VOD)	– Offering unlimited service for a one flat monthly or yearly fee	Netflix, Watcha
AVOD (Advertising VOD)	– Offering free content based on exposure to advertisements	YouTube, hulu
FVOD (Free VOD)	– Offering free content available to anyone – Used to expand subscriber base, etc.	New platforms, existing platforms
Hybrid VOD	– A combination of at least two models among TVOD, SVOD, AVOD, and FVOD	YouTube, oksusu, olleh TV Mobile, U + Mobile TV

earlier in the study, the OTT service based on broadcasting contents offered mainly by telecom operators, at the initial stages of a service launch, set up a business model based on broadcasting content in order to secure subscribers. Yet, with the exception of several big-name producers, it's difficult for CPs to earn revenue due to a gap in bargaining power with broadcasters. Broadcasters make an advance payment as a minimum guarantee (MG) to take over copyrights for a big portion of content; and such payments make up for a shortfall in production costs resulting from the poor management conditions of CPs. Even if OTT operators purchase these contents and constitute a lineup of contents, the revenues earned therefrom will go to broadcasters, rather than CPs. Even in the case of old contents, it's difficult to expect any revenue therefrom because ole contents are used to set up an archive on the cheap.

Recently, terrestrial broadcasters' contents are increasingly losing influence. In 2015, terrestrial broadcasters' real-time channel service was interrupted due to disputes over fees for re-transmitting terrestrial broadcasting content. At the time, some IPTV users defected to other services, but soon, the user base returned to the previous years' levels, and subscriber numbers continued to increase. As noted earlier, it is safe to say that in the past, the success of the launch of new media was heavily dependent on terrestrial broadcasters' contents supply and demand. Now that formula for success has changed. Also, with contents of paid broadcasting business operators, such as CJ ENM, gaining popularity, reliance on terrestrial broadcasters' contents is declining even further. What matters now is what we see, instead of through what

medium. That means producing original content becomes crucial to differentiate oneself from other operators.

Recently, Telecom OTT services, which are based on broadcasting contents, have begun making concentrated efforts to create their own content. In 2018, SK Broadband produced and serviced the original content, entitled "I Picked Up A Star On The Road" through its OTT service 'oksusu'. 'oksusu' covered 100% of the production cost, and it succeeded, garnering 7 million views for two months. In addition, "I Picked Up A Star On The Road" was sold to Japan, Taiwan, and other countries even before the launch of broadcasting service, leading to generous profits.

If such new attempt is made in cooperation with domestic small-and medium-sized independent media producers, it will be really helpful for strengthening the self-production capacity of OTT operators as well as the ecosystem of independent broadcasting companies. However, given that sharing benefits serves as a great incentive for stakeholders to co-exist and thrive, the rationality of new contracts and revenue distribution standards should be guaranteed as they launch new services.

3.3.3 Original Content Type

The original content type is to build a self-production lineup as described above. Of course, not all service archives are created internally. It also serves contents and movies from existing broadcasters. Yet, it aims to lower dependence on existing content, such as terrestrial broadcasting contents and movies strategically. It will also strengthen its own production capabilities, expand its influence as a new platform, and create a new content ecosystem. 'Netflix' is a prime example of a self-produced original content. From the revenue sharing (benefit sharing) viewpoint, the original-content type seems more viable and reasonable than the broadcasting content type for several reasons. First of all, original-content type OTT providers invest in CPs to produce content on their own. From CPs' perspective, these investments are really helpful for them to produce content and to manage their operations. Investment made by OTT providers make up shortfalls in the production costs of small-and-medium-sized CPs, an arrangement similar to the MG contract, under which existing broadcasters pay advance payment to CPs. The original-content-based OTT operators fully cover production costs as an investment, and in addition, pay approximately 20% more as business income and operational expenses. That translates into stable production costs as well as a certain amount of revenue for CPs. That's why CPs seek to expand into new business. In addition, establishing partnerships with operators with global presence, such as Netflix, can be instrumental in marketing their works overseas. That means addressing both issues—raising production costs and marketing the content—at the same time.

'Studio Dragon,' an independent producer and production affiliate of the Korean entertainment giant CJ ENM, signed a license agreement with Netflix for broadcasting rights of Mr. Sunshine, a 24-episode series. Through the deal, the producer raised 30 billion won out of some 40 billion won in total production costs. The downside is that the RS contract, which would guarantee additional revenue if the content takes

off, was not adopted. As Netflix is also operating based on the copyright agreement, the revenues generated through the content usage go to OTT operators.

4 Conclusion

Recently, it is increasingly becoming an essential and important source of competitive advantage to establish and utilize desirable inter-organization relationships in the Korean broadcasting content industry. Such trend, along with uncertainty fundamental and inherent in the broadcasting content, will only grow stronger down the line amid market uncertainty in the Korean content industry.

This study has shed light on the inter-organizational revenue sharing as a way to enhance competitiveness in terms of co-evolution through inter-organizational relations in the Korean broadcast contents industry. Amidst the murky Korean broadcast content environment, several conditions challenging independent competitiveness enhancement are calling for cooperative co-evolution in terms of inter-organizational relationship; and competitiveness co-evolution in terms of industrial relations. This study has attempted an exploratory approach to the ways of strategically sharing benefits which can deepen and sustain motivation and effort for strengthening the mutual inter-organizational relationships from a cooperative co-evolutionary perspective.

In particular, it is necessary to look for ways of strategic benefit sharing between OTT, which is a new distributor, and CPs, in addition to the relationship between traditional TV broadcasters and CPs. Amid competition between TV and OTT, sharing benefits with CPs is getting more and more attention. In this study, we attempted an exploratory typology based on the circumstances in which TVs and OTT share the benefits with CPs on the basis of distribution channels that have opposite but complementary features, to some extent.

As mentioned earlier, such competition among businesses is taking place through various types of contracts among an array of operators. This study focused on two major issues that are central to benefit sharing (revenue sharing): A traditional issue of distribution dealing with revenue sharing between terrestrial broadcasters and CPs; and the newly emerging issue of distribution centering around OTT operators. No major change has occurred in the traditional way of distribution; While a new ecosystem is being created with diversification of the media, the country still relies heavily on terrestrial distribution of contents.

However, the issue of distribution centering around OTT operators shows the possibility of creating a new media ecosystem and establishing more reasonable benefit sharing (revenue sharing) standards. OTT operators pay enough production costs and operating expenses to CPs. This indicates opportunities for CPs, who have had relatively weak bargaining power, to expand new business and to create a new ecosystem. OTT operators, however, are also adopting a business model focused on securing copyrights. In other words, most of the revenue generated after the production is completed goes to OTT operators. This shows limitations in terms of continuous benefit sharing (revenue sharing). Of course, there is no need to guarantee contin-

uous benefit-sharing for all goods, products, and services. One-off transactions can also be concluded. Yet, the problem lies in the fact that other than selling copyrights, chances of sharing other benefits (revenue distribution) are slim. In addition, as OTT operator-oriented service type is expanded to original content production, numerous attempts and considerations need to be given with regard to how to share benefits.[5]

Such discussion process takes a more forward-looking perspective in that it not only expands the scope of discussion beyond benefit sharing between TV broadcasters and CPs to include benefit sharing between OTT and CPs, but also in that this process deals with strategic benefit sharing for economic co-evolution, not just the one for social co-evolution. In addition, such process takes on new meaning in that rather than defining specific ways of benefit sharing as the most desirable one, it explores various ways of desirable benefit sharing depending on the situation each partner is in.

In the future, it will be necessary to verify actual benefits earned through sharing the benefits of the inter-organizational relationship, which can be called prototype of some types, and to execute and verify the feasibility of the type presented in this exploratory study with no clear prototypes or examples. In particular, In particular, although the typology is based on the Korean broadcasting contents industry, what we need is to conduct an in-depth analysis and take supplementary measures to identify the commonalities and differences of the Korean broadcasting content industry, and those of advanced economies in the global context. By doing so, it will be possible to establish specialized ways of benefit sharing for the Korean broadcasting content industry.

This study has several limitations. In addition, these limitations present necessary research. Firstly, The phenomenon of OTT affecting the broadcasting content industry is universal and benefit sharing (revenue sharing) is of common importance, but can it be generalized through the situation of Korean broadcasting content industry.

Secondly, it looked at benefit sharing as a way of enhancing competitiveness from the perspective of sustainable co-evolution. Yet, even if these types of benefit sharing have, in a sense, strategic and rational aspects, further verification is needed to see if they will fuel co-evolution which can lead to sustainable, long-term relations in the relational context, not just the short-term desirable relations in the transactional context. In particular, a more macroscopic approach is also needed to look at competitive co-evolution that takes place in relations between each organization from an industrial point of view, not just bilateral relations.

Thirdly, the study excluded benefit sharing centering around one-person media, such as MCN, in exploratory typology of diversification of distribution channels, an issue mentioned at the beginning. This may be done for the purpose of parsimony to allow for more clear comparison between TV and OTT, but more consideration needs to be given thereto down the line at a time when individualized production systems and broadcasting contents are increasingly becoming commonplace.

[5]Currently, OTT service can be classified into broadcasting content-type and original content-type. Yet, as shown in the main body, original content production is becoming more and more important, rendering any attempt to the classification of service types pointless.

Fourthly, although the study took into account different situations in its exploratory approach for typology, it failed to give consideration to internal factors unique among broadcasters, OTT and CPs (broadcasters (terrestrial or comprehensive programming channels, cables, etc.), characteristics of OTT, the types, size, history/reference, etc. of CPs, etc.) This can be a critical factor in making the typology more sophisticated. In particular, a failure to take into account strategic direction each player is seeking can be viewed as one of the limitations of the study.

Lastly, despite an imbalance between contribution and inducement up to a certain point and a dynamic equilibrium in a snapshot, co-evolution, like a seesaw which eventually reaches equilibrium in the long run, can work effectively from a relatively long-term perspective. Given that, the study is also limited in that it fails to propose policy, institutional tools for ensuring a long-term balance. This can be a task that needs to be tackled in the future work.

Acknowledgements "This research was supported by the Academic Research fund of Hoseo University in 2017" (2017-0071).

References

1. Korea Creative Contect Agency: <2017 Broadcasting industry white paper> (2018)
2. Noh, M.S.: A study on the utilization status and policy tasks for SMEs' Sharing System. Wage Information Brief **2**(3), 11–15 (2018)
3. Lee, S.H.: Current situation and prospect of broadcasting contents in Korea towards the changing circumstances of broadcasting market. Focus **28**(8), 1–17 (2016)
4. Baek, S.H., Byun, S.K., Kim, G.J.: Broadcasting contents market and profit distribution model. Korean J. Broadcast. Telecommun. Stud. **29**(1), 63–92 (2015)
5. Baek, S.H., Kim, G.J.: An exploratory study on the fair trade and profit structure of broadcasting contents. Korean J. Broadcasting Telecommun. Stud. (2013). (Regular academic conferences in spring, in Busan)
6. Korea Communications Commission.: The competitive Broadcasting Market Situation Assessment (2017)
7. Jeon, B.S.: OTT service at the center of the global media market. Broadcast. Trend Insight **14** (2018)
8. Kim, Y.K., Nam, S.H.: An Study on Films in the Digital Online Market Structure Analysis. Korean film Council (2016)

A Study of Utilizing Backend as a Service (BaaS) Space for Mobile Applications

Cheong Ghil Kim

Abstract With the advancement of semiconductor and information and communication technologies, our computing environments are evolving into complex, large-scale, and high-performance platforms. Especially, this movement is characterized as hyper-connected society, in which IoT and cloud computing are core technologies. The Parse Server is a full feature, production ready Parse Platform on the Cloud. It could be an optimal platform for IoT service by providing open-source back-end services for developers with a set of tools to enable front-end development focus. In this paper, Parse Server is interoperable with an open source based game and a single board computer to develop use cases for the purpose of IoT programming education. This research aims to get insights of methodology for IoT programming with implementation of game ranking service and remote sensing systems. Implementation results were confirmed with Parse Dashboard and Rasberry Pi.

Keywords Cloud computing · Internet of thing · Parser server · Open source · Backend as a service

1 Introduction

Nowadays, it is understood that we are going into the hyper-connected society in which people and people, people and objects, things and things, online and offline, one to one, one to many and many-to-many are connected using digital technology or the digital based society that stores, analyzes, and manages data collected by sensors and mobile terminal through a network [1, 2]. The commencement of 5G commercialization services may accelerate this movement; people and objects can be interconnected intelligently.

Figure 1 shows the conceptual prediction might be aroused by starting with 5G networks made by EU project METIS which stands for Mobile and wireless communications. They are characterized as avalanche of traffic volume, massive growth

C. G. Kim (✉)
Department of Computer Science, Namseoul University, 91 Daehak-ro Seonghwan-eup, Cheonan, South Korea
e-mail: cgkim@nsu.ac.kr

R. Lee (ed.), *Computer and Information Science*, Studies in Computational Intelligence 849, https://doi.org/10.1007/978-3-030-25213-7_15

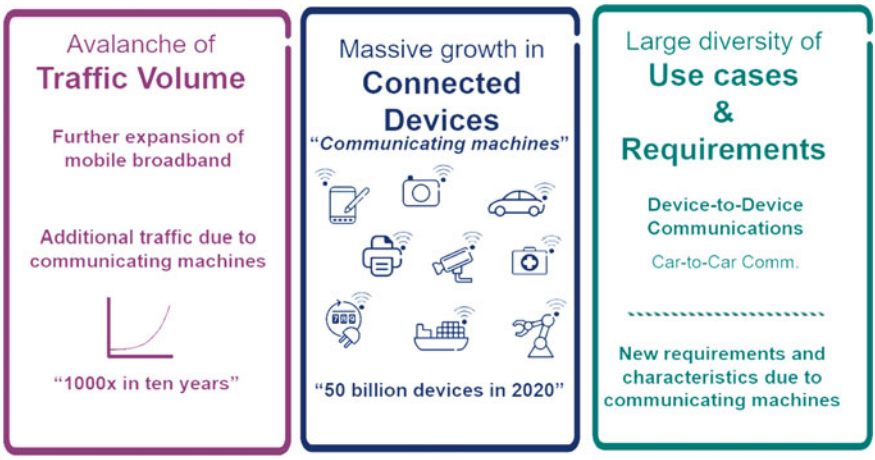


Fig. 1 Conceptual prediction with 5G

in connected devices, and large diversity of use cases and requirements. Especially, 50 billion communication devices are expected to be connected. Along with this process, IoT (Internet of Thing) has already become a part of many industries, including the automotive, construction and medical industries. In general, the development of the IoT has been primarily driven by needs of large corporations that stand to benefit greatly from the foresight and predictability afforded by the ability to follow all objects through the commodity chains in which they are embedded [3]. The ability to code and track objects has allowed companies to become more efficient, speed up processes, reduce error, prevent theft, and incorporate complex and flexible organizational systems through IoT [4]. The IoT is a technological revolution that represents the future of computing and communications, and its development depends on dynamic technical innovation in a number of important fields, from wireless sensors to nanotechnology. They are going tag each object for identifying, automating, monitoring and controlling [5].

In addition, cloud services are becoming more diverse, evolving and innovating across society, going beyond collective intelligence based on information exchanged between objects through IoT (Internet of Things) [12, 13].

Figure 2 shows a service block diagram of M2 M which communicates between machine and machine via network and all the networks are connected [9]. This may be the basic technology for IoT which refers to the evolution of existing USN (Ubiquitous Sensor Network). While M2M was mainly aimed at communication with end-devices and people, IoT extends the scope of things so that people can communicate with various sensors and devices around us. Furthermore, IoT technology can be advanced as the object-space communication network that forms intelligent relationships such as sensing, networking, and information processing in a cooperative way without human intervention to the distributed environment elements.

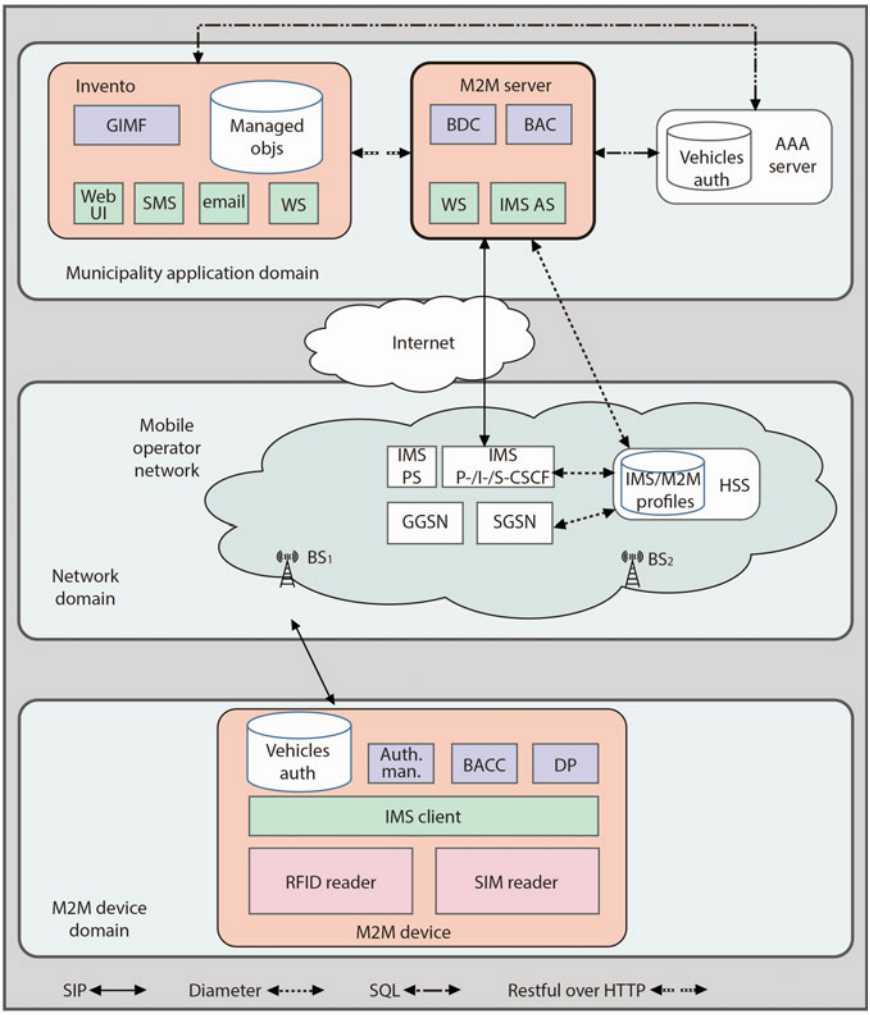


Fig. 2 A distributed M2M service architecture

The concept of cloud computing addresses the next evolutionary step of distributed computing. The goal of this computing model is to make a better use of distributed resources, put them together in order to achieve higher throughput, and be able to tackle large scale computation problems. Cloud Computing is not a completely new concept for the development and operation of web applications. It allows for the most cost-effective development of scalable web portals on highly available and fail-safe infrastructures [6].

With the importance of information being emphasized, Facebook acquired Parse service platform and it was subsequently open sourced allowing individual developers to handle hardware and software at no cost [4, 5] while adding Arduino Yun

and Embedded C SDK to BaaS (Backend as a Service) which only supported mobile application development.

Parse Server [4] provides a back-end, application development platform that provides programmers with a set of tools to focus on front-end development; currently it is used in many applications. As an open-source platform, it is understood that Parse provides one of the easiest ways to get a database and RESTful API up and running. So, in case that users want to build a mobile app and don't want to code the back-end by hand, Parse could be a good solution. Figure 3 shows a way of building data driven applications with Parse. Users can deploy their own Parse data store and push notifications systems to Heroku leveraging the server open-sourced by Parse. Once the Parse server is configured, users can initialize Parse within Android app pointing the client to their self-hosted URL. After that, the functions demonstrated in this guide work the same as they did before. Parse is built on top of the MongoDB database which can be added to Heroku using MongoLab [7].

Therefore, in this paper, Parse Server is interoperable with an open source based game and a single board computer to develop use cases for the purpose of IoT programming education verifying the possibility of IoT system programming education. This research aims to get insights of methodology for IoT programming with implementation of game ranking service and remote sensing systems. Implementation results were confirmed with Parse Dashboard and Rasberry Pi. The composition of this paper is as follows. In the next section, we introduces the types and parse of cloud services. In Sect. 3, the game ranking system is designed by linking Parse Server. In Chap. 4, the implementation result is introduced through Parse Dashboard. We conclude the last chapter.

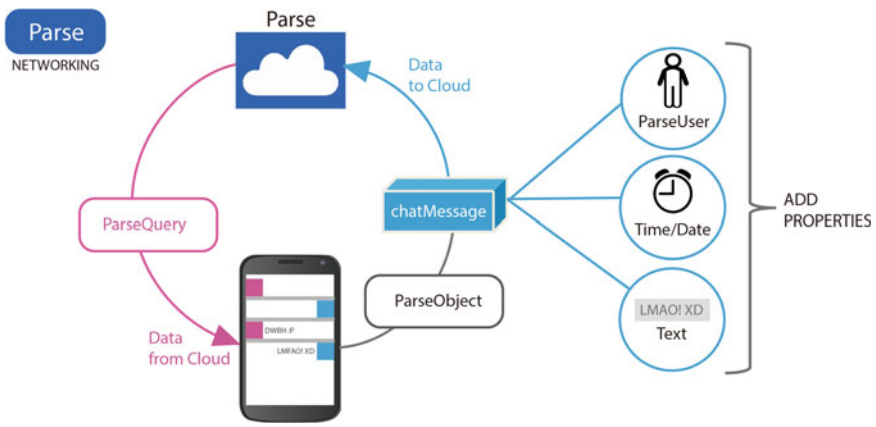


Fig. 3 Data driven apps with Parse

2 Background

2.1 Cloud Architecture

The value of data in the Fourth Industrial Revolution is becoming more emphasized. Clouds are spreading as a foundation to provide various software and services remotely from anywhere, and to store and manage huge amounts of data in the process [6]. The concept of cloud computing addresses the next evolutionary step of distributed computing. The goal of this computing model is to make a better use of distributed resources, put them together in order to achieve higher throughput and be able to tackle large scale computation problems [10, 11].

Figure 4 shows the categorized cloud architecture according to their service offerings. In general, there are three categories of cloud services such as infrastructure, platform, application etc. These services are delivered and consumed in real-time over the Internet.

(1) Software as a Service (SaaS): Software as a Service is a multi-tenant platform. It uses common resources and a single instance of both the object code of an application as well as the underlying database to support multiple customers simultaneously. Sometime, it is referred to as the Application Service Provider (ASP) model; and it is a service provided directly to the user such as Dropbox or Google docs.

(2) Platform as a Service (PaaS): Platform as a Service is a service that provides a platform. Therefore, it can provide developers with a platform including all the systems and environments comprising the end-to-end life cycle of developing, testing, deploying and hosting of sophisticated web applications as a service

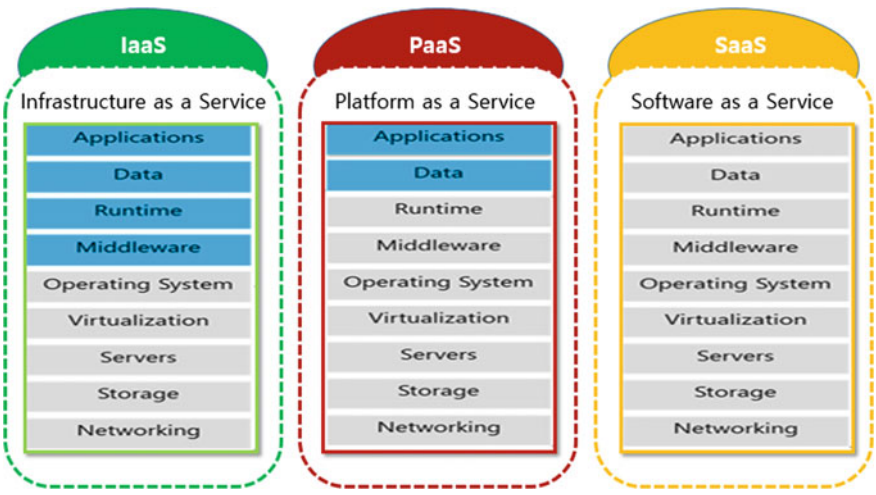


Fig. 4 Cloud architecture based on services offering

delivered by a cloud. Compared with conventional application development, this architecture can same development time, offer hundreds of readily available tools and services, and quickly scale. Both web server and middleware could be a representative; Microsoft's Azure is a key example.

(3) Infrastructure as a Service (IaaS): Infrastructure as a Service is the delivery of computer infrastructure as a service. Aside from the higher flexibility, a key benefit of IaaS is the usage-based payment scheme. This allows customers to pay as they grow. Another important advantage is that by always using the latest technology. Customers can achieve a much faster service delivery and time to market. Traditional hosting service is a good example.

In this figure, the services provided by the cloud are shown in gray color; blue is the area that the user must deal with. Infrastructure as a Service (IaaS) is a service that provides IT infrastructure resources such as servers and storage over the Internet. Platform as a Service (PaaS) is a service that provides the platform needed to develop software services. The user develops the application by selecting the required services in PaaS. Software as a Service (SaaS) is an application service that runs in a cloud environment, all services are in the cloud.

In addition to that, there is an emerging trend in mobile application development, that is called backend as a service (BaaS), this is a cloud service for data only in a recent classification. The main reason developers need a server when building a mobile application is to send and receive data. Figure 5 shows the architecture of Baas which concentrates on the communications between devices. Signing up and logging in user accounts, storing data in applications, collecting logs, and so on, are all done at the data exchange level. In order to send and receive data, a suitable server framework should be selected, server software should be developed based on the server framework, and the server software should be deployed so that it can be accessed through the server hosting or web hosting. However, BaaS provides a service that solves the difficult task of combining devices and software [6].

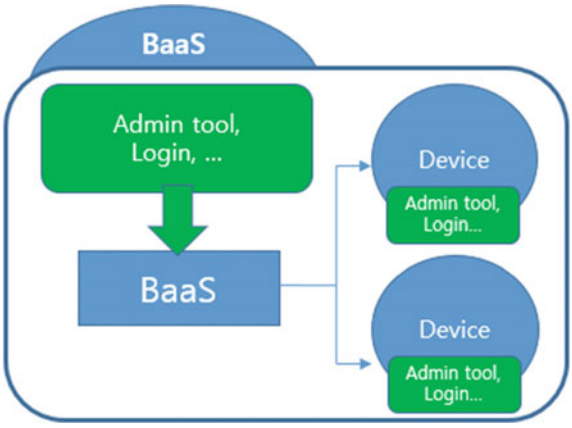


Fig. 5 Architecture of backend as a service (BaaS)

2.2 *Parse*

Modern applications resort to specialized backend storage systems to store their persistent state, such as user accounts, shared content, documents and purchases. Replication, not only across various server replicas but also on the mobile device itself, is widely adopted by these backends in order to improve the reliability and performance. Although maintaining the strong consistency among replicas can guarantee the correctness of application behaviors, it will affect the application performance at the same time because there is a well-known trade-off between consistency and performance [14].

Because Parse is a cloud service, you can follow the classification criteria of the cloud, and by that point it is BaaS (Backend as a Service). To briefly describe this, Parse can be defined as a cloud for exchanging data only. Parse was a cloud services platform that supported the early mobile backend. The addition of Arduino Yun and the Embedded C SDK has enabled us to leverage the IoT system. Figure 7 is an SDK supported by Parse. Since then, Parse has been turned into open-source software that can be installed, reducing the time and cost individuals spend on server configurations. Since the function of sending and receiving data is already registered in Parse, the configuration of the server is completed by registering the application on the web page providing the solution and registering the application.

As shown in Fig. 6 the Parse server allows data to be exchanged between devices, Web, and apps created by individuals. Storing the data from the sensors in the Parse Server can be used on the web or in the app. Conversely, when you save data sent from the web or an app, you can read the values and control the devices accordingly (Fig. 7).

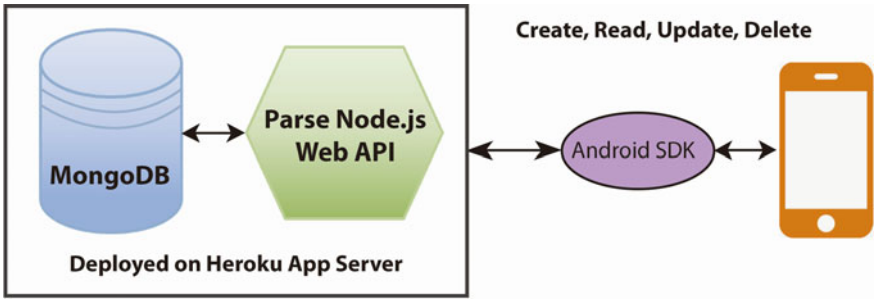


Fig. 6 Overview of web services with Heroku App server

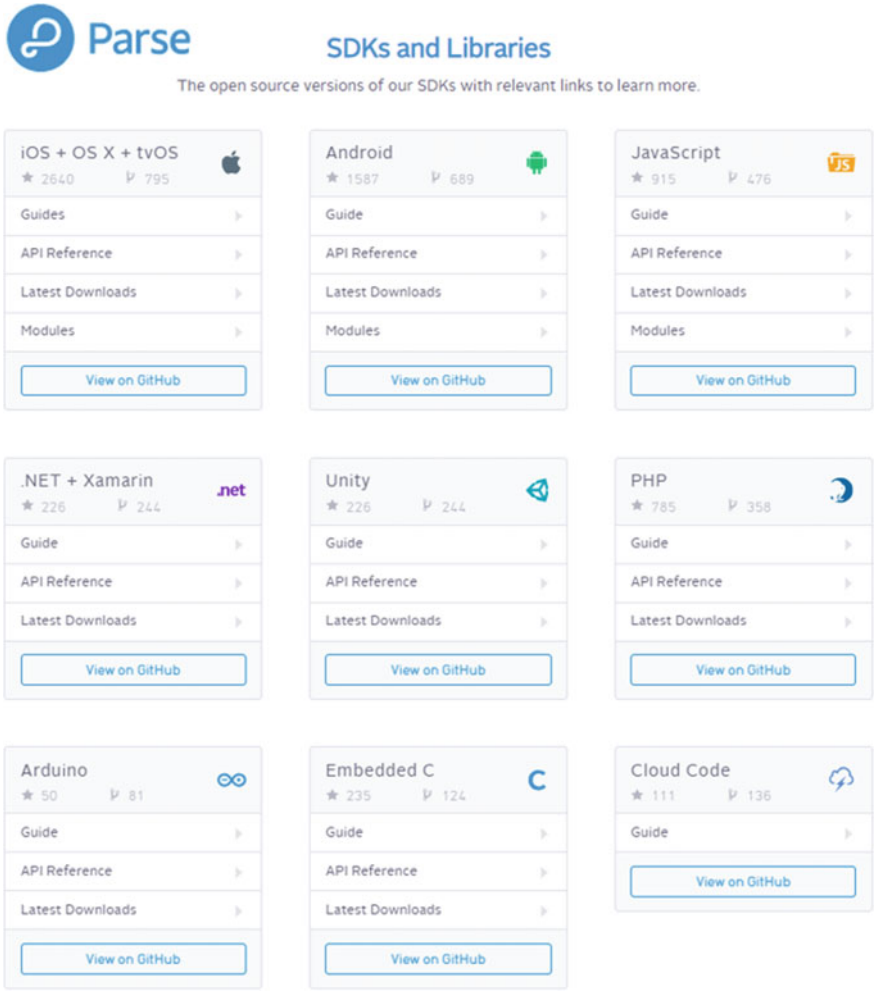


Fig. 7 SDK supported by Parse

3 Implementations

This section introduces an implementation of a system that provides ranking information by storing game score points and using stored data in conjunction with a Parse service that can exchange data. Figure 8 shows the overall system architecture implemented.

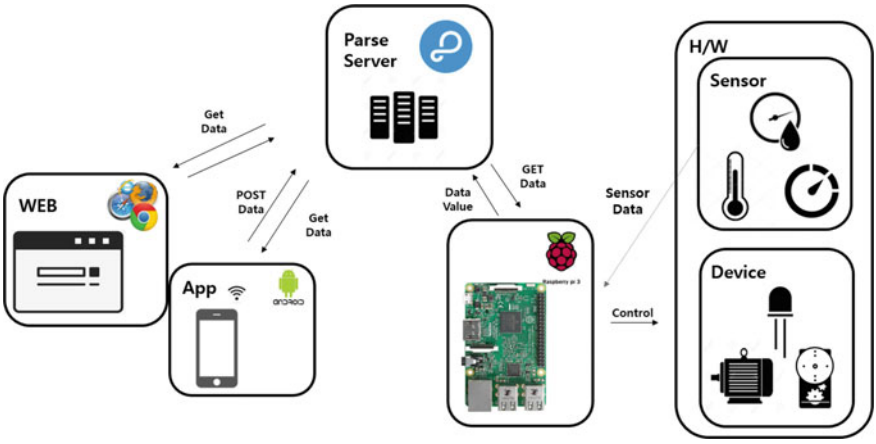


Fig. 8 Overall system architecture using Parse server

3.1 Game Ranking Service

Even simple games, data exists in all games. One of the basic elements in the game is the scoring system. Scoring system is an important tool to encourage users to play with their sense of accomplishment and competition with other users. However, most open-source Web games are only one-off games because they do not save, only output at the end of the scoring system [8].

Parse Server is a 'card matching game', an open source game based on the JavaScript language. Arrange the cards randomly, tell the location of the card in a short time, and flip it over to the same colored face. The game user must remember the location of the same colored card in a short period of time and match all cards. Outputs the play time when matching all cards. If you hit OK, the record you played is a one-off game that disappears. This data is stored and utilized by Parse Server to build a game ranking system shown in Fig. 9.

3.2 Remote Sensing and Controlling

In Fig. 10, We implement a system that performs remote control of LED and transmission save of occurrence event using Raspberry Pie. To do this, we created a Web program using Javascript. As a result, the *html* file which is responsible for the actual operation and the *php* file which is necessary for web distribution are generated. The remote device was implemented by controlling the GPIO port of Raspberry using the *wiringPi* library. After that, The Web distribution module using Heroku was implemented.

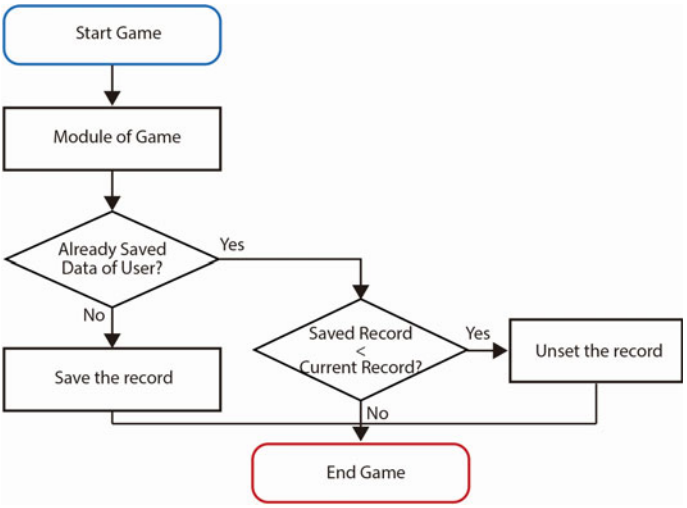


Fig. 9 Processing flow of game ranking service using Parse server

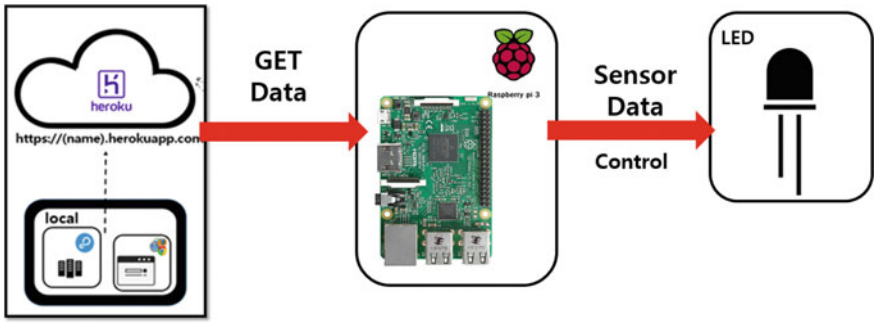


Fig. 10 Remote LED control with Parse server

4 Results

Figure 11 shows the process comparison between the processes with and without Baas. Taking advantage of BaaS, we can seve the development time by excluding server developing time. Figure 12 is the result of users playing the game and storing the resulting values in Parse Server. The stored values can be checked through the Parse Dashboard [7, 14]. Using the stored data, it is the result of the implemented ranking system. In a one-time game, we built a game system that supports ranking service by storing and utilizing data after server construction. Figure 13 shows the implementation result in which users can switch on and off on the web and communicate with a remote Raspberry Pi; all the events are stored in the Parse server which can be verified with the Parse Dashboard.

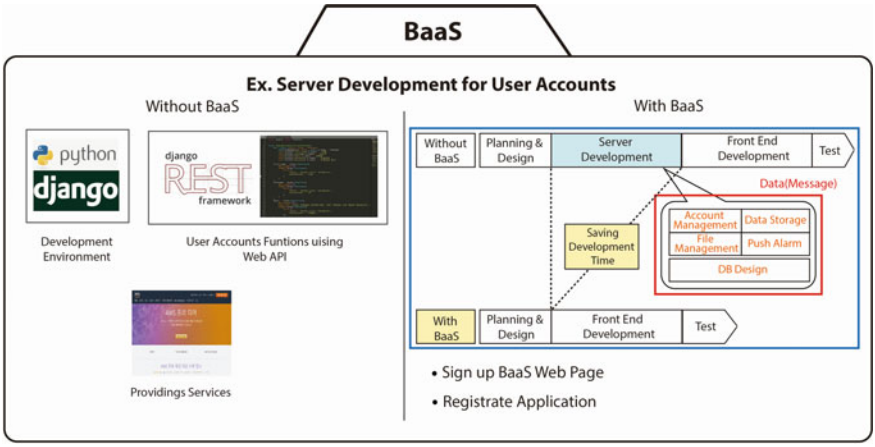


Fig. 11 Process comparison using Parse server

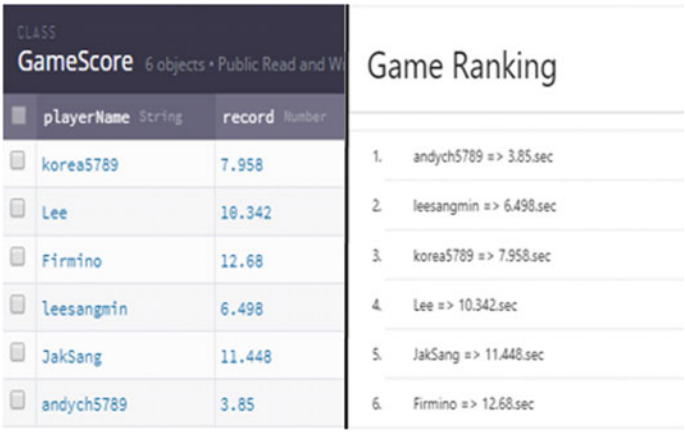


Fig. 12 Ranking system result

Fig. 13 Remote sensing and control system result

5 Conclusion

The Parse Server is a full features, production ready Parse Platform on the Cloud. It could be an optimal platform for IoT service by providing open-source back-end services for developers with a set of tools to enable front-end development focus. In this paper, Parse Server is interoperable with an open source based game and a single board computer to develop use cases for the purpose of IoT programming education. This research aims to get insights of methodology for IoT programming with implementation of game ranking service and remote sensing systems. Two systems were implemented effective by saving server development process and their implementation results were confirmed with Parse Dashboard.

Acknowledgements Fundings for this paper was provided by Namseoul University.

References

1. Chung, S.W., Park, S.J., Lee, S.Y.: The era of hyper-connected society and the challenges in business activities: focusing on information blocking and acquisition activies. Int. J. Manage. Appl. Sci. **3**(4), 16–21 (2017)
2. Park, J., Yoon, M.Y.: Hyper connected society and future services. J. Korean Inst. Commun. Sci. (Inf. Commun.) **31**(4), 3–9 (2014)
3. Lianos, M., Douglas, M.: Dangerization and the end of deviance: the institutional environment. Br. J. Criminol. **40**, 261–278 (2000)
4. Ferguson, T.: Have your objects call my object. Harv. Bus. Rev. 1–7 (2002)
5. Madakam, S., Ramaswamy, R., Tripathi, S.: Internet of things (IoT): a literature review. J. Comput. Commun. **3**, 164–173 (2015)
6. Parse Server.: [Online]. Available: https://parseplatform.org/
7. The AWS and MongoDB Infrastructure of Parse: Lessons Learned: [Online]. Available: https://medium.baqend.com/parse-is-gone-a-few-secrets-about-their-infrastructure-91b3ab2fcf71
8. Lee, S.M., Lee, Y.S., Kim, C.G.: A design and implementation of game ranking service system using parse server platform. Proc. KIPS Spring Conf. **26**(1) (2019)
9. Ray, P. P.: A survey on internet of things architectures. J. King Saud Univ. Comput. Inf. Sci. **30**, 291–319 (2018)
10. Grønli, T., Hansen J., Ghinea G.: Android vs Windows Mobile vs Java ME: a comparative study of mobile development environments. In: Proceedings of the 3rd International Conference on Pervasive Technologies Related to Assistive Environments, ACM, New York, 1–8 (2010)
11. Rimal, B.P., Choi, E., Lumb, I.: A taxonomy and survey of cloud computing systems. In: Proceedings of the 2009 Fifth International Joint Conference on INC, IMS and IDC, pp. 44–51 (2009)
12. Kim, C.G.: Web of Things: future internet technologies and applications. J. Platf. Technol. **1**(1), 33–40 (2013)
13. Kim, C.G., Lee, S.H.: A study on web of things technologies for smart building energy management system. In: Proceedings of the 3rd International Conference on Convergence Technology, pp. 318–319 (2013)
14. Tavares, A.: Variable data consistency in Parse Server, [Online]. Available: https://fenix.tecnico.ulisboa.pt/downloadFile/1407770020546078/extended-abstract-Variable-data-consistency-in-ParseServer-76395.pdf

Author Index

R. Lee (ed.), *Computer and Information Science*, Studies in Computational Intelligence 849, https://doi.org/10.1007/978-3-030-25213-7