

IWIN2017



International Workshop on Informatics

Proceedings of
International Workshop on Informatics

September 3-6, 2017

Zagreb, Croatia



Sponsored by Informatics Society

IWIN2017



International Workshop on Informatics

Proceedings of
International Workshop on Informatics

September 3-6, 2017

Zagreb, Croatia



Sponsored by Informatics Society

Publication office:

Informatics Laboratory

3-41, Tsujimachi, Kitaku, Nagoya 462-0032, Japan

Publisher:

Tadanori Mizuno, President of Informatics Society

ISBN:

978-4-902523-41-6

Printed in Japan

Table of Contents

Session 1: IoT System and Devices

(Chair: Yoshia Saito) (9:10 - 10:10, Sep. 4)

- (1) Development of Management System for the office and the PC practice room using IoT technology 3
Yuki Takabayashi, Yu Kagami, Haruka Kawauchi, Masao Isshiki, and Keiichi Abe
- (2) Development of a Scheduler with Visualizing Electricity Consumption and Controlling Home Appliance 9
Takumi Shida, Shigeaki Miyazawa, Hiroshi Sugimura, Keiichi Abe, and Masao Isshiki
- (3) Building Smart Block and Base Recognizing 3D Electric Circuit 13
Masaki Ohtani, Kazuki Hayashi, Kazuki Tokunaga, Junnosuke Takeuchi, Norikane Kanai, and Masao Isshiki
- (4) Acceleration Method for Real-time Processing in Distributed Control System 19
Akihiro Yamashita, Hiroshi Mieno, and Tadanori Mizuno

Session 2: Intelligent Transportation Systems

(Chair: Takuya Yoshihiro) (10:40 - 12:00, Sep. 4)

- (5) A Traffic Simulation for Improving Traffic Flow in Provincial Cities.....31
Chinatsu Ikeda, Masashi Saito, and Ryozo Kiyohara
- (6) A Study of Driver's Contexts for Visibility of Equipment on Vehicle.....39
Yuki Chigira, Yuichi Tokunaga, and Ryozo Kiyohara
- (7) VIS Performance Analysis in Multi-Lane T-Junction49
Hiroki Ito, and Yusuke Takatori
- (8) Evaluation of Embedded Platform for Autonomous Drive Using Linux and GPU.....53
Shota Mizuno, Shinya Honda, Ryo Kurachi, Daisuke Hayakawa, and Naoya Chujo

Keynote Speech 1

(13:30 - 14:10, Sep. 4)

- (I) Digital Forensics in Japan.....61
Prof. Ryoichi Sasaki, Tokyo Denki University

Session 3: Data Models

(Chair: Tomoki Yoshihisa) (14:20 - 15:40, Sep. 4)

- (9) Non-Model Approach based Process Improvement Method 123
Akihiro Hayashi, and Nobuhiro Kataoka
- (10) Derivation of a Map of Variables in a Loop Structure..... 129
Kozo Okano, Shinji Kusumoto, and Yukihiro Sasaki
- (11) A Method for Verifying Equivalence of Functions in C language..... 135
Satoshi Harauchi, Kozo Okano, and Shinpei Ogata
- (12) A Rule-based Method of Stepwise Evaluating Class Diagrams..... 141
Shinpei Ogata, Kazune Miyajima, Mizue Kayama, and Kozo Okano

Session 4: Multimedia Systems

(Chair: Katsuhiko Kaji) (15:50 - 17:10, Sep. 4)

- (13) TinyASIO: A C++ Library for using ASIO with Short Source Code 149
Eiichi Takebuchi, Tomoki Kajinami, Ichiro Tokuhiko, and Haruo Hayami
- (14) A Method for Video Advertisement Insertion with Audience Comments on Action Game
Videos 153
Yoshia Saito
- (15) WakWak Tube: An Asynchronous YouTube Co-Viewing System..... 159
Yuichi Bannai, and Yusuke Wakatsuki
- (16) A Design and Implementation of Distributed Internet Live Broadcasting Systems
Enhanced by Cloud Computing Services..... 165
Satoru Matsumoto, Yoshimasa Ishi, Tomoki Yoshihisa, Tomoya Kawakami, and
Yuuichi Teranishi

Session 5: Position Recognition

(Chair: Yu Enokibori) (9:00 - 10:20, Sep. 5)

- (17) A Method for Estimating Indoor Position and Walking Speed by Using a Smartphone, with the Aim of Grasping People Flow in Indoor Passages..... 173
Masaru Matsubayashi, and Yoh Shiraishi
- (18) A Proposal of Method to Correct GPS Positioning Error by using Acceleration Sensor 183
Chiaki Yashiro, and Tomoyuki Yashiro
- (19) A BLE Beacon's Movement and Equipment Defect Estimation Method Based on Comparison of Room-level Wi-Fi and BLE Fingerprints..... 189
Shota Ikeda, Fumitaka Naruse, and Katsuhiko Kaji
- (20) A Prototype of a Precision Forecasting System for Real-Time Navigation with RTK-GNSS 193
Tiphath Areeyapinun, and Tomoya Kitani

Session 6: Systems and Applications

(Chair: Kozo Okano) (10:30 – 12:10, Sep. 5)

- (21) Statistical Prediction of Wagyu Characteristics with Bayesian Hierarchical Modeling · 203
Hisami Sano, Takatoshi Fujiki, Takuya Yoshihiro, Haruka Ikegami,
Tamako Matsuhashi, and Kazuya Matsumoto
- (22) Best-time Estimation Method by Region and Tourist Spot using Information
Interpolation 209
Masaki Endo, Masaharu Hirota, Shigeyoshi Ohno, and Hiroshi Ishikawa
- (23) Automatic Recording of Actors' Body Expressions with Action Cueing 217
Takayoshi Takano, Hiroshi Shigeno, and Ken-ichi Okada
- (24) Discovering Hotspots Using Photographic Orientation and Angle of View from Social
Media Site 223
Masaharu Hirota, Masaki Endo, Daiju Kato, and Hiroshi Ishikawa
- (25) An Approach for E-garment based Skill Analysis with Accelerometer based Skill Segment
Extraction 229
Yu Enokibori, Shinya Namikawa, and Kenji Mase

Keynote Speech 2

(13:30 - 14:10, Sep. 5)

- (II) Learning with Gamification, Collaboration and Augmented Reality Technology..... 239
Prof. Ivica Botički, Faculty of Electrical Engineering and Computing,
University of Zagreb

Session 7: Network and Security

(Chair: Yoh Shiraishi) (14:20 - 15:40, Sep. 5)

- (26) A Scheduling Algorithm for Slotted-CSMA Based Wireless Mesh Networks with High-speed Links 249
Yasuhiro Mori, and Takuya Yoshihiro
- (27) A Strategy for Mitigating Information Loss in Anonymization by using Nature of Personal Data 255
Takayasu Yamaguchi, and Hiroshi Yoshiura
- (28) A Classification Method of Unknown Malicious Websites Using Address Features of each Network Address Class 261
Shihori Kanazawa, Yoshitaka Nakamura, Hiroshi Inamura, and Osamu Takahashi
- (29) Online Magnitude Fluctuation Analysis for Anomaly Detection in Equipment Condition Monitoring 269
Makoto Imamura, Daniel Nikovski, Amir-massoud Farahmand, and Harumi Watanabe

Session 8: Industrial Applications

(Chair: Ryozo Kiyohara) (15:50 - 17:10, Sep. 5)

- (30) A Study on Optimum Mega-Photovoltaic System Design 279
Hironori Honda, Yasushi Tachibana, Jin Umeoka, Masashi Saito, and
Yasumitsu Suzuki
- (31) A Trial of Ambient Sensor Method for Worker States Classifier 289
Kazuyuki Iso, Minoru Kobayashi, and Takaya Yuizono
- (32) Smart Robot based on Web Resource with Friendly Communication and Its Application to
Library Guide 295
Yoichi Yamazaki, Takuya Shinkai, and Masao Isshiki
- (33) A Comparative Study in Enterprise Systems Dealing with Images: MongoDB vs. MySQL
..... 301
Tsukasa Kudo, and Yuki Furukawa

Message from the General Chairs



It is our great pleasure to welcome all of you to Zagreb, Croatia, for the Eleventh International Workshop on Informatics (IWIN 2017). This workshop has been held annually by the Informatics Society. Since 2007, the workshops were held in Naples in Italy, Wien in Austria, Hawaii in the USA, Edinburgh in Scotland, Venice in Italy, Chamonix in France, Stockholm in Sweden, Prague in Czech Republic, Amsterdam in Netherlands, and Riga in Latvia respectively.

In IWIN 2017, 33 papers were accepted after peer reviewing by the program committee. Based on the papers, eight technical sessions were organized in a single track format, which highlighted the latest research results in the areas such as Internet of Things (IoT), mobile computing, networking, data models, multimedia systems, geographical information systems. IWIN 2017 will also welcome two keynote speakers: Prof. Ryoichi Sasaki of Tokyo Denki University, Japan and Prof. Ivica Botički of University of Zagreb, Croatia. We really appreciate their participation in the workshop.

We would like to thank all the participants and contributors who made the workshop possible. It is indeed an honor to work with a large group of professionals around the world for making the workshop a great success. We are looking forward to seeing you all in the workshop. We hope you enjoy IWIN 2017.

September 2017

Ken-ichi Okada
Tomoo Inoue

Organizing Committee

General Chair

Ken-ichi Okada (Keio University, Japan)
Tomoo Inoue (University of Tsukuba, Japan)

Steering Committee

Toru Hasegawa (Osaka University, Japan)
Teruo Higashino (Osaka University, Japan)
Tadanori Mizuno (Aichi Institute of Technology, Japan)
Jun Munemori (Wakayama University, Japan)
Yuko Murayama (Tsuda College, Japan)
Ken-ichi Okada (Keio University, Japan)
Norio Shiratori (Chuo University / Tohoku University, Japan)
Osamu Takahashi (Future University Hakodate, Japan)

Program Chair

Tomoki Yoshihisa (Osaka University, Japan)

Financial Chair

Tomoya Kitani (Shizuoka University, Japan)

Publicity Chair

Yoshitaka Nakamura (Future University Hakodate, Japan)

Program Committee

Keiichi Abe (Kanagawa Institute of Technology, Japan)	Yuichi Bannai (Kanagawa Institute of Technology, Japan)
Naoya Chujo (Aichi Institute of Technology, Japan)	Chiaki Doi (NTT DOCOMO, Inc., Japan)
Yu Enokibori (Nagoya University, Japan)	Teruyuki Hasegawa (KDDI R&D Laboratories, Japan)
Takaaki Hishida (Aichi Institute of Technology, Japan)	Hiroshi Horikawa (Mitsubishi Electric Information Network Corporation, Japan)
Makoto Imamura (Tokai University, Japan)	Hiroshi Ishikawa (Tokyo Metropolitan University, Japan)
Hiroshi Inamura (Future University Hakodate, Japan)	Masaji Katagiri (NTT DOCOMO, Inc., Japan)
Katsuhiko Kaji (Aichi Institute of Technology, Japan)	Tomoya Kitani (Shizuoka University, Japan)
Yoshinobu Kawabe (Aichi Institute of Technology, Japan)	Minoru Kobayashi (Meiji University, Japan)
Ryozo Kiyohara (Kanagawa Institute of Technology, Japan)	Tsukasa Kudo (Shizuoka Institute of Science and Technology, Japan)
Yuki Koizumi (Osaka University, Japan)	Shinichiro Mori (Chiba Institute of Technology, Japan)
Fusako Kusunoki (Tama Art University)	Yoshitoshi Murata (Iwate Prefectural University, Japan)
Hiroshi Mineno (Shizuoka University, Japan)	Yoshitaka Nakamura (Future University Hakodate, Japan)
Hiroaki Morino (Shibaura Institute of Technology, Japan)	Ken Ohta (NTT DOCOMO, Inc., Japan)
Katsuhiro Naito (Aichi Institute of Technology, Japan)	Yoshia Saito (Iwate Prefectural University, Japan)
Tetsushi Ohki (Shizuoka University, Japan)	Tetsuya Shigeyasu (Prefectural University of Hiroshima, Japan)
Kozo Okano (Shinshu University, Japan)	Yusuke Takatori (Kanagawa Institute of Technology, Japan)
Fumiaki Sato (Toho University, Japan)	Hirozumi Yamaguchi (Osaka University, Japan)
Yoh Shiraishi (Future University Hakodate, Japan)	Shohei Yokoyama (Shizuoka University, Japan)
Takaaki Umedu (Shiga University, Japan)	Tomoki Yoshihisa (Osaka University, Japan)
Tomoyuki Yashiro (Chiba Institute of Technology, Japan)	Takaya Yuizonon (Japan Advanced Institute of Science and Technology, Japan)
Takuya Yoshihiro (Wakayama University, Japan)	
Hiroshi Yoshiura (The University of Electro-Communications, Japan)	

Session 1:
IoT System and Devices
(Chair: Yoshia Saito)

Development of PC Management System for the office and the PC practice room using IoT technology

Yuki Takabayashi[†], Yu Kagami[‡], Haruka Kawauchi[‡], Masao Isshiki^{*}, and Keiichi Abe[‡],

^{†, ‡, *}, Kanagawa Institute of Technology, Japan

[†]takabayashi2017@ele.kanagawa-it.ac.jp, [‡]{s1433018,s1433009,abe}@he.kanagawa-it.ac.jp,

^{*}Masao@kait.jp

Abstract - At present time, there have been a demand for management systems that can survey and monitor a personal computer (PC) practice room, movement of people in an office, situation of utilization of facilities and so on in real time without causing psychological stress. For example, PC administrators in Japanese national educational institutions must report operation status of PC practice rooms once a year. But, there is currently no system for automatically recording PC operating situations. Therefore, the burden on the PC administrators is big. In this study, we aimed at systems for accurately managing the sitting and work time without psychologically stressing PC users. This time, we propose uniform management systems of sitting and work time using smart tap node and mat sensor node by IoT technology. The smart tap was connected to PCs to acquire the operating status of the PCs. In parallel with this smart tap, the mat sensor was used to acquire human presence state. By calculating binary data representing the PC operating status and human presence state from the two sensing data (smart tap and mat sensor) by the proposed technique, we can exactly calculate wasteful power consumption etc. The use of IoT technology makes it unnecessary to use large installation services when introducing our system. Therefore, this our proposal system can be easily installed even by an amateur.

Keywords: Management System, Sitting and Work time, Smart tap node, Mat sensor node, IoT technology

1 INTRODUCTION

There is a demand for a system which can survey and monitor the use situation of facilities and movements of persons in a personal computer (PC) practice room or office in real time without psychologically stressing the users. For example, an annual report of the operation situation (the number of users, use time) of PC practice rooms is required in some Japanese educational institutions. This is considerably burdensome for system managers because the tasks of recording, counting, and reporting the operation situation typically rely on manual work, with poor utilization of IT tools.

In recent years, universities and other educational institutions use PCs frequently regardless of the field. PCs have also been essential in businesses as well as in educational institutions. With the number of users and devices increased to dozens, or even hundreds, the burden on managers has been rapidly increasing.

Therefore, in this study, we propose a management system for administration in a PC practice room or office and the PC operation situation.

The system described in this paper utilizes IoT (Internet of Things) technology by wireless communication (Wi-Fi). This system requires no major construction and can be easily installed even by an amateur.

This paper is organized as follows. Chapter 2 describes related technologies. Chapter 3 gives an overview of the proposed system, and Chapter 4 details the prototype development. In Chapter 5, we will describe the evaluation results of the prototype. In Chapter 6, this research is summarized.

2 RELATED WORKS (OR STUDY)

Examples of the conventional techniques for checking the presence and attendance of users in a PC practice room include an attendance system which checks the entry and leaving times by IC cards, and an attendance management system which determines attendance of users in a lecture based on a database which holds PC use histories [1] [2] [3] [4].

There are also studies on automated systems for checking the entry and leaving times in a room, such as a system which employs the iBeacon technology, a hands-free system for checking the entry, presence leaving of users, and an attendance checking system which uses a camera [5] [6] [7] [8] [9].

Those foregoing studies enable us to grasp the entering and leaving times of users, but there is the problem that we cannot exactly know the actual work time. So, we consider a system for exactly managing the attendance time of users and the PC work time. Using a camera would certainly enable an exact, real-time monitoring of the work time. However, constant monitoring by camera may make some users feel psychological stressed.

Therefore, in this study, we propose a system for real-time administration of the attendance time of users in a PC training room or office and the PC work time.

Unlike the conventional studies which require big installation work for system introduction, our system requires no such big work and can be installed even by amateurs.

3 SUMMARY OF THE SYSTEM

In this chapter, we describe the management system using the smart tap node and mat sensor node.

Figure 1 shows the outline of the proposed system for managing the operating status of each personal computer in a PC practice room or office.

In this system, in order to manage the operation status of each personal computer, the number of PC users is calculated from the presence status of the users and the PC operation time is obtained from the PC work situation. Then, we devised a system which can total the operating situation over the year and finally automatically submit it as a report.

The proposal of this research aims to realize a management system using smart taps and mat sensors as described.

This system used the IoT technology by wireless communication. We used a wireless communication standard Wi-Fi (2.4 GHz) as the interface between wireless nodes such as smart tap and mat sensor and host computer for data collection. For this reason, this system requires no major construction and can be easily installed even by an amateur.

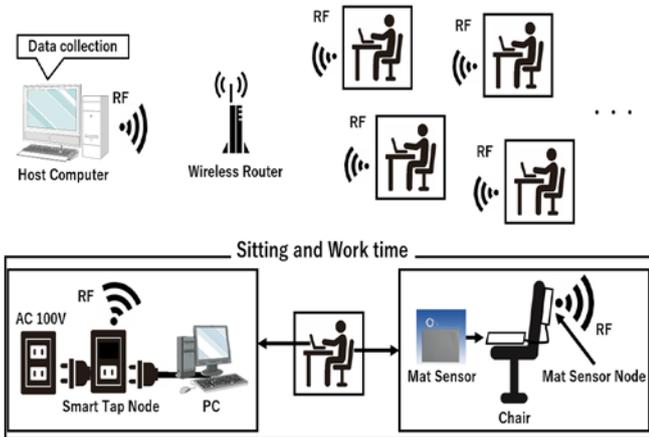


Figure 1: Overview of Our system.

3.1 Proposal of Calculation Method of Setting and Work Time

We install a smart tap for the PC as a method to check the operation situation of the PC. We can find ON/OFF of the PC power supply from the power consumption information acquired with the smart tap. We can estimate the work time of the PC user from this ON/OFF information to some extent. However, the estimation may become inaccurate depending on the action of the user. For example, suppose that a user has moved to another place with the PC power on. Since the PC power supply is ON, our system will mistakenly determine that a user is working at the PC.

Therefore, we installed a mat sensor as a method to detect the presence of a user sitting in the chair. This system allows for more exact detection of the presence of a user.

For the calculation of the PC work time, we totaled the periods of time as “PC work time” when there was a user in the chair and the PC power supply was ON. But this may incorrectly include the time when the user sitting in the chair is doing some non-PC work while the PC power is ON. So,

we additionally tested the technique of changing the period of the sleep time for PC to more exactly determine the PC work time.

Based on these ideas, we devised a system for exactly managing the user attendance and PC work time from two kinds of data acquired with the mat sensor installed in the chair and the smart tap installed for the PC.

From the two kinds of sensor data acquired with the smart tap and mat sensor, we can estimate the activity of the user in four ways as shown in Table 1.

At first we understand ON/OFF of the PC from power consumption information of smart tap. "1" of smart tap shows that the PC is ON, and "0" shows that the PC is OFF. In the system, the power consumption threshold is 15 W. Therefore, if it exceeds 15 W, the output is "1" and less than 15 W is "0". On the other hand, "1" of the mat sensor indicates a state where someone is in the chair, and "0" indicates no one is in the chair. It is possible to manage one PC by the two pieces of sensor information.

When A = 0 and B = 0, the PC is in OFF state and no user is present, hence "Absence". A = 0, B = 1 indicates that the PC is in the OFF state, but because someone is present, it indicates "Work time other than PC". When A = 1 and B = 0, the PC is in the ON state, but because no one is present, it indicates "Wasted electric use time". When A = 1 and B = 1, since the user is present while the PC is ON, this state is judged to be "PC work time" in this study.

Table 1: Each sensor information and action estimate of the person

A Smart tap (PC-ON/OFF)	B Mat sensor (Sitting/Not sitting)	Action estimate of the person
0	0	Absence (PC-OFF/Not sitting)
0	1	Work time other than PC (PC-OFF/Sitting)
1	0	Wasted electric use time (PC-ON/Not sitting)
1	1	PC work time (PC-ON/Sitting)

With output “W” indicating the absence of a user, “X” indicating someone doing non-PC work, Y indicating the PC wastefully consuming electricity, and “Z” indicating someone working on PC, the following equations (3.1) to (3.4) can be derived from Table 1. The output of the smart tap is A, and the output of the mat sensor is B.

$$W = \bar{A} \cdot \bar{B} \quad (3.1)$$

$$X = \bar{A} \cdot B \quad (3.2)$$

$$Y = A \cdot \bar{B} \quad (3.3)$$

$$Z = A \cdot B \quad (3.4)$$

Although we defined the output Z of equation (3.4) as PC work, it is not always the case that the user sitting in the chair is actually doing PC work; the user may possibly be doing different work. So, in order to distinguish between PC

work and separate work, we conceived the idea of actively utilizing PC sleep function in addition to the two kinds of sensor data. We conducted an experiment to evaluate this idea.

4 IMPLEMENTATION

Figure 2 shows an overview of our system per one person's by smart tap node and mat sensor node.

In the system proposed in this study, one smart tap node and one mat sensor node are allocated to one subject and used, to confirm the operation status of one person's PC.

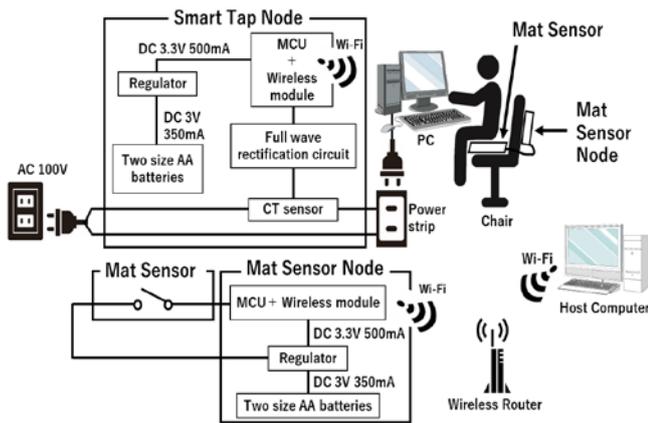


Figure 2: Overview of our system by smart tap node and mat sensor node per one person's.

4.1 Smart Tap Node

The smart tap node with wireless communication developed this time calculates the total value of current consumption in CT sensor (KM20-CTF-50A/Omron). Power (apparent power) is simply calculated from the product of the voltage that is constant at AC 100V and the measured current. Power consumption information is delivered to the host computer via Wi-Fi. Commercially available MCU + wireless module (ESP-WROOM-02/Espressif Systems) was used for Wi-Fi communication [10]. It has Micro Controller Unit (MCU) built in and can be used as Arduino. The Wring language, which is a C-based language, was used for the software development this time.

4.2 Mat Sensor Node

In this system, the presence or absence of a user sitting in a chair was detected by using a mat sensor (M-A4/Takenaka Engineering). As for the Wi-Fi communication, a wireless module with built-in MCU was used like the smart tap node. Figure 3 shows a photograph with mat sensor node mounted.

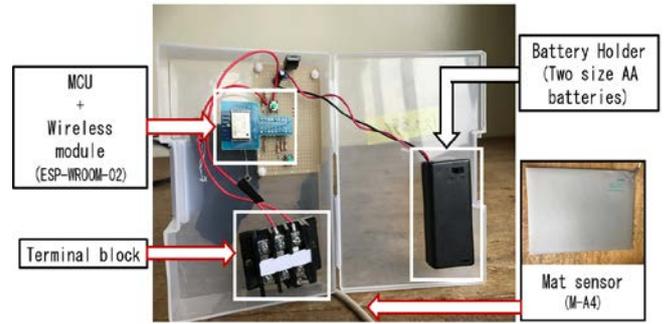


Figure 3: Overview of mat sensor node.

4.3 Firmware of MCU with Wireless Node

Figure 4 shows flowchart of data transmission firmware for MCU (Type: ESP-WROOM-02) with wireless node.

The wireless module used for each node was built as an Arduino with MCU built in.

The wireless module with built-in MCU operates to transmit the measurement result to the host computer by Wi-Fi communication.

To allow for data transmission from multiple nodes, the IP address and port number of the host computer are set in advance in the MCU so that the destination can be specified. In this study, UDP/IP communication was used as the communication method.

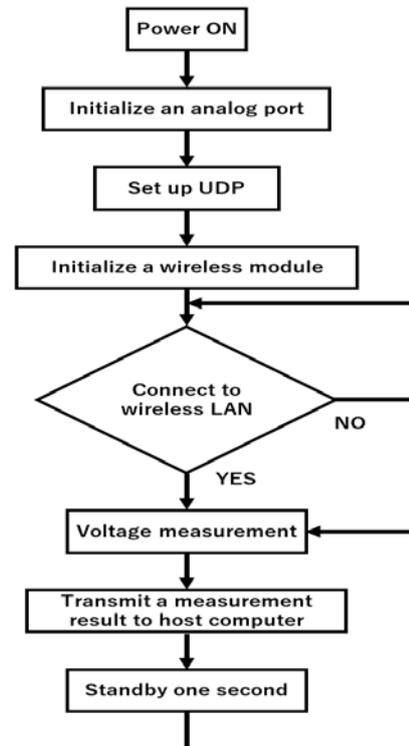


Figure 4: Flowchart of Data Transmission Firmware for MCU with Wireless Node.

4.4 Application Software of Host Computer

On the host computer, we developed an application that can display and store the collected data in real time (Figure 5).

This application calculates PC work time and wasted electric use time from the two pieces of information including the person's presence status and PC work situation, and displays them on the screen. Work time other than PC and absence time are also calculated and displayed on the PC screen.

This Graphical User Interface (GUI) application was developed in Visual Basic 2015 development environment.

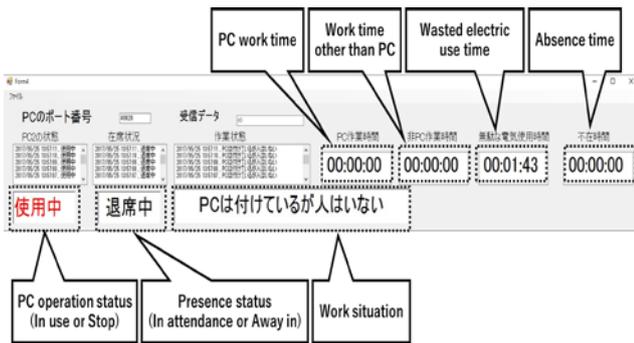


Figure 5: GUI application on host computer.

5 EVALUATION OF THE SYSTEM

In this Chapter 5, we evaluate the prototype system we developed.

5.1 Experimental Environment

Figure 6 shows a layout of the smart tap node and mat sensor node by our system. The empirical evaluation of the prototype was carried out at the Kanagawa Institute of Technology C2 building 6th floor the Room No. E602 (ABE-Laboratory). The evaluation period is from May 8, 2017 to May 26, 2017. In this study, the PC operation status and PC work situation for two subjects were measured.

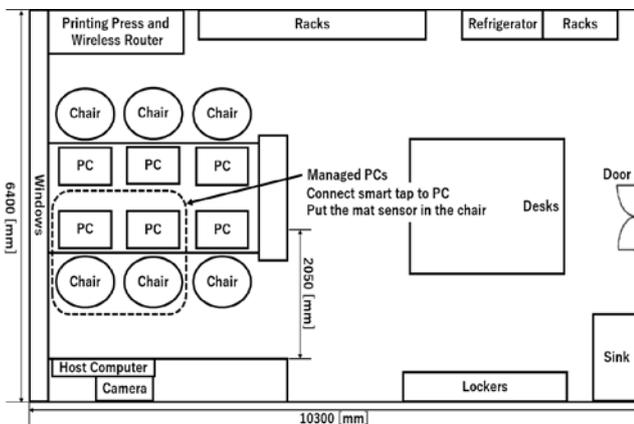


Figure 6: Layout of smart tap node and mat sensor node by our system.

This time, one host computer was used for data management. Smart taps and mat sensors were installed on each of the two PCs. Figure 7 shows actual experiment image.

A real-time video was also simultaneously acquired and used as the “correct answer” data. For the real time display, we used the application created in Visual Basic 2015.

Actual work time etc. were obtained from the correct answer data and compared with aggregated data in prototype application for evaluation. No user was at PC 1 (left) while a user was present at PC 2 (right) (Figure 7). The moving image from the web camera connected to the host compute were taken in real time as correct answer data.

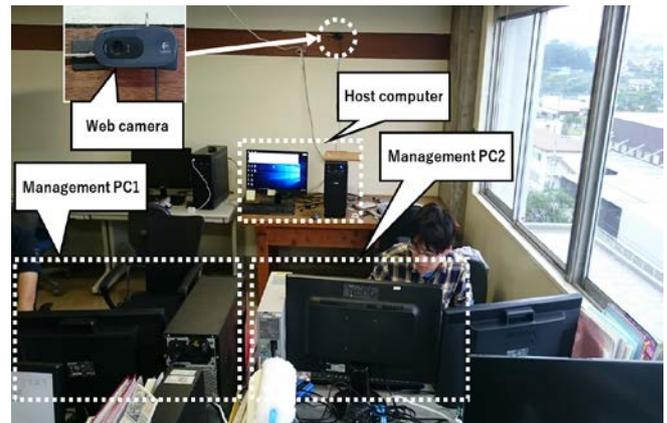


Figure 7: Experiment image.

5.2 Comparative Evaluation of Camera Data VS Our Prototype System

In the proposed method, there are four states (“PC work time”, “work time other than PC”, “wasted electric use time”, and “absence time”) for which the period of time can be calculated. The time of each of these four states is calculated by the application on the host computer.

For the initial evaluation, we compared the degree of accuracy of the prototype system with the moving images shot with the camera.

The time values of the four states acquired by the prototype system are shown in Table 2 and Figure 8.

The result demonstrated that the error between was minus one second in both wasted electric use time and absence time, suggesting that those measurements were highly accurate. The total work time (PC work time and working time other time PC) was also found to be fairly accurate, with a minor error of minus 9 seconds. In contrast, there is a large error of about 1 hour in PC work time and work time other than PC. In the prototype system, the "work time other than PC" must be in the state in which the mat sensor is ON and the PC power is OFF. The large error seems to have occurred because the period of time when the user was performing work other than PC with the power of the PC turned on was erroneously recorded as “PC work time” by the prototype system.

Table 2: Comparative evaluation of camera(video) vs our prototype system

	PC work time	Work time other than PC	Total work time	Wasted electric use time	Absence time
Camera (video)	0:48:46	1:06:03	1:54:49	0:11:26	0:00:54
Prototype System	1:54:41	0:00:17	1:54:58	0:11:25	0:00:53
Error	+1:05:55	-1:05:46	+0:00:09	-0:00:01	-0:00:01

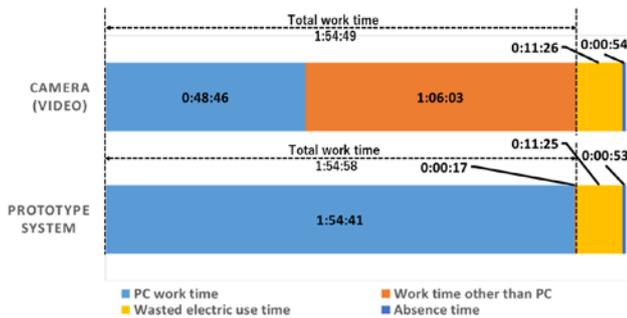


Figure 8: Comparison between camera image and prototype (not sleep mode).

5.3 Dependence on PC Sleep Time

As mentioned in the previous section, the period of time when PC work is performed with PC power on is recorded as "PC work time". Therefore, we thought that we could more accurately calculate PC work time and work time other than PC by using PC sleep mode. The interval of the sleep time was 1 minute and 10 minutes. Since the PC's sleep mode time setting was at least 1 minute, 1 minute was used in this experiment.

Experimental results show that the PC work time calculated by the prototype system and the work time other than the PC are closer to the camera time when the sleep time interval is one minute than when 10 minute. The error is also smaller in 1 minute than in 10 minutes. Namely, "PC work time" and "work time other than PC" can be accurately obtained if the PC sleep time interval setting is short. Setting the sleep time at 1 minute rather than 10 minutes also reduces the wasted electric use time by 10 minutes or even more.

It was confirmed that the prototype system can more accurately measure PC work time etc. by using the sleep mode of PC. Also, using the sleep mode decreases wasted electric use time, so power saving effect can be expected.

Table 3: Sleep time interval 10 minutes

	PC work time	Work time other than PC	Total work time	Wasted electric use time	Absence time
Camera (video)	1:34:24	0:25:19	1:59:43	0:14:43	0:01:37
Prototype System	1:50:31	0:09:14	1:59:44	0:14:45	0:01:34
Error	+0:16:07	-0:16:05	+0:00:01	+0:00:02	-0:00:03

Table 4: Sleep time interval 1 minutes

	PC work time	Work time other than PC	Total work time	Wasted electric use time	Absence time
Camera (video)	2:22:45	0:17:10	2:39:55	0:02:24	0:05:07
Prototype System	2:25:57	0:14:47	2:40:44	0:02:41	0:04:55
Error	+0:03:12	-0:02:23	+0:00:49	-0:00:03	-0:00:12

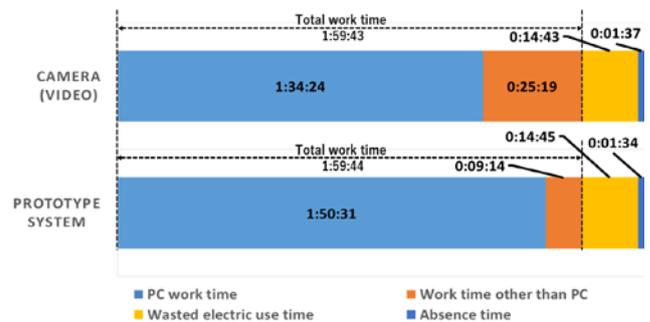


Figure 9: Result when sleep time is set to 10 minute.

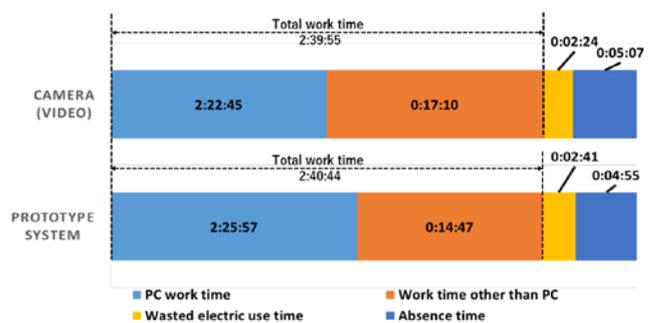


Figure 10: Result when sleep time is set to 1 minute.

6 CONCLUSIONS AND FUTURE WORK

In this study, we proposed a method to calculate the four kinds of time (PC work time, work time other than PC, absence time, wasted electric use time) from data acquired from smart tap and mat sensor. We also developed a prototype that implemented this proposed method and conducted a demonstration experiment. As a result, it was possible to accurately obtain the total time of the PC work and the work other than the PC, the absence time, and the wasted electric use time. Evaluation experiments revealed that "PC work time" and "work time other than PC" depend on the sleep time of PC. The influence of the PC sleep time on "PC work time" and "work time other than PC" could be decreased by setting a short PC sleep time. The minimum sleep time of the PC is 1 minute, and there is a limit to actually use it. Also, if the sleep time is too short, PC users may feel stressed. In the future, we would like to consider a method that can accurately monitor PC work time without stressing PC users.

In this evaluation experiment, we conducted with two examinees, but we would like to conduct an evaluation experiment in a large number of people (PC practice room and offices).

REFERENCES

- [1] N. Morita, "Development and Result of Attendance Registration System Using IC cards (In Japanese)", The Institute of Electronics, Information and Communication Engineers (IEICE)/Forum on Information Technology (FIT), 11(3), pp.65-68, (2012).
- [2] S. Kubota, S. Furukawa, Y. Soejima, R. Kawamura, and K. Sugitani, "presence-type Attendance Management System in Educational PC Classroom (In Japanese)", The Institute of Electronics, Information and Communication Engineers (IEICE)/Forum on Information Technology (FIT), 8(4), pp.175-177, (2009).
- [3] H. Suda, S. Nakamura, M. Ogawa, H. Kumazawa, N. Komdou, "Attendance Management System using Password Distribution of NFC Tag with Electronic Papers (In Japanese)", The 76TH National Convention of IPSJ, No.4, pp.853-854, (2014).
- [4] H. Matsumoto, S. Wada, S. Hara, N. Morita, "Development and Evaluation of Attendance Registration System using Suicas (In Japanese)", Proceedings of the School of Information and Telecommunication Engineering Tokai University, Vol.1, No.2, pp.16-21, (2011).
- [5] K. Tanaka, K. Shuwa, "Application of iBeacon in Laboratory Room Management System Automations (In Japanese)", Tokyo City University Yokohama Campus Journal of Information Studies, Vol.4, No.16, pp.31-37, (2015).
- [6] N. Takayama, M. Kitamura, "Hand Free Entry/Attendance Management Systems (In Japanese)", NEC Technical Report, Vol.63, No.3, pp.60-63, (2010).
- [7] S. Ono, H. Yasuda, "Proposal of Setting Management System", Proceedings of the 2014 IEICE General Conference, pp.184, (2014).
- [8] K. Onozawa, K. Watanabe, K. Suwa, "Room Management System and Learning Situation Management System Using Smartphones (In Japanese)", Tokyo City University Yokohama Campus Journal of Information Studies, Vol.4, No.13, pp.6-15, (2012).
- [9] H. Suda, S. Nakamura, M. Ogawa, H. Kumazawa, N. Komdou, "Attendance Management System using Password Distribution of NFC Tag with Electronic Papers (In Japanese)", The 76TH National Convention of IPSJ, No.4, pp.853-854, (2014).
- [9] N. Nishimura, "New Suddenly Attends (In Japanese)", Journal of Yasuda Women's University, Vol.41, pp.371-379, (2013).
- [10] CQ publishing companies, "Web brain connection Wi-Fi × 3G/LTE of IoTs (In Japanese)", Transistor technology, Vol.9, pp.35-93, (2016).

Development of a Scheduler with Visualizing Electricity Consumption and Controlling Home Appliance

Takumi Shida*, Shigeaki Miyazawa*, Hiroshi Sugimura*, Keiichi Abe*, and Masao Isshiki*

*Graduate School of Engineering, Kanagawa Institute of Technology, Japan

{shida2017, miyazawa2016}@ele.kanagawa-it.ac.jp,

{sugimura, abe, masao}@he.kanagawa-it.ac.jp

Abstract -Home Energy Management System (HEMS) has been drawing attention with the development of schedulers for controlling home appliances and visualizing electricity consumption as well as the spread of IoT home appliances. The HEMS promotes electricity-saving habits by visualizing electricity consumption at home and controlling home appliances. However, visualization of electricity consumption is merely aimed at promoting users' awareness of energy saving and is not intended to be continuously used. For continuous use of this function, it is necessary to add useful functions for users. We propose a scheduler with the functions of controlling home appliances and visualizing electricity consumption. When displaying a home appliance controller, the scheduler places icons within a graph which visualizes electricity consumption. The simultaneous display of the two functions makes users inevitably view the graph on a continuous basis. The system also supports electricity-saving habits by displaying control icons for home appliances as well as showing the schedule of deenergizing home appliances. In this study, we have developed a tool for promoting energy saving and evaluated it by questionnaires.

Keywords: Visualizing electricity consumption, Scheduler, Controlling home appliances, Electricity saving, HEMS

1 INTRODUCTION

With the spread of IoT home appliances, the Home Energy Management System (HEMS) has been drawing attention. The HEMS promotes electricity-saving habits by controlling home appliances and visualizing electricity consumption. Its electricity-saving effect has been proven by a test, in which the electricity consumption in a home decreased by approximately 17 % over three months [1]. However, it has also been revealed that the energy-saving effect declines as the use period becomes longer. In study [2], a survey was conducted on the use of the visualizing function by families who tried out the HEMS for about three months. The survey showed that the number of users who used the HEMS every day decreased to one half over the three-month period. This result suggests that some measures for supporting the continuous use of the system are needed to obtain a long-term electricity-saving effect. To solve this problem, various studies for providing the HEMS with additional values have been conducted.

Study [3] proposes a system for proposing the most suitable advice on energy saving based on power consumption and peak hours. The system selects the best

advice for users based on advice lists. However, in some cases, this method gives users a useless piece of advice.

Study [4] proposes an energy-saving method by an automatic control of home appliances. However, if there are two or more individuals in the same room, it is difficult to create a comfortable environment by the automatic control.

Study [5] proposes several energy-saving plans for the entire system of the home appliances. By executing the presented plan, users can promote an energy-saving lifestyle without sacrificing comfort. However, users need to have empirical knowledge on power consumption to specify a desired target value for energy savings.

Study [6] proposes visualization of the power consumption of home appliances using a label whose size indicates the amount of power consumed in one day. Users are required to put the label at a point of high electricity consumption on a graph. Such a task hinders continuous use of the HEMS.

We point out that a continuous use is essential for the HEMS. For this purpose, we propose a scheduler that has the two functions of controlling home appliances and visualizing electricity consumption.

The scheduler displays a controller when an icon representing a home appliance is placed on a graph area which visualizes electricity consumption. The simultaneous display of the two functions makes the user inevitably view the electricity consumption visualized on the graph. When the user controls a home appliance, the scheduler prompts the user to set the power-on period of the home appliance. In this paper, we will describe the details of the scheduler and evaluate it by using questionnaires.

2 SUMMARY OF THE SYSTEM

Figure 1 shows the principal idea on the user interface for the scheduler. The scheduler has the functions of visualizing current consumption and controlling home appliances. The two functions are shown on the screen. If many appliances are connected in a house, the scheduler will be cluttered with many items displayed. To avoid this, the scheduler displays icons of home appliances and their product names in its main area. The icons allow users to easily operate home appliances. The scheduler displays a graph of the current consumption and the icons in the main area, as well as a sub area for displaying a controller below the main area. Furthermore, two functions are added to the scheduler.

The first function is to check the operations of each home appliance. The second function is to set power-on periods

for those appliances. We will develop a system that runs on a web browser. The system should meet three specifications:

- (1) Visualization of current consumption
- (2) Control of home appliances
- (3) Visualization of the operating state of home appliances

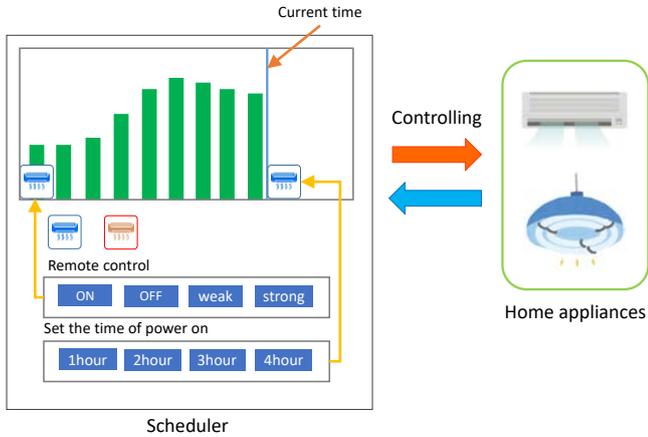


Figure 1: The principal idea on user interface for the scheduler.

2.1 Visualizing electricity consumption

Switchboards for domestic use have a specific amount of maximum current. Since exceeding the specified amount causes a blackout, users need to reduce the amount of current consumption. For this purpose, the scheduler visually displays the amount of current consumption and improves the users' awareness of electricity saving. The scheduler retrieves 24-hour data from the switchboard. It requires no measuring instruments to be attached to the individual home appliances to obtain their respective data.

2.2 Controlling of home appliances

The home appliance controller separately displays a screen for controlling home appliances and one for programming their power-on periods. Users can display the two types of screens by dragging icons. When an icon is dropped on the graph, a controller screen, which corresponds to the icon as well as its position on the graph, is displayed. The controller is displayed below the graph displayed on the screen. Users can control the home appliances while checking the graph.

2.3 Visualizing state of home appliances.

When an icon is dropped in the scheduler, the icon is placed on the graph. The icon changes its shape depending on the operation time and current consumption of the corresponding home appliance. The icons allow users to visually check the on/off state and current consumption of each appliance. The scheduler includes the function of stacking a new icon on top of an already existing one. With this function, users can comprehend the entire consumption current in the home as well as that of each home appliance.

3 IMPLEMENTATION

Figure 2 shows an entire system configuration. The data used for the visualization of current consumption was obtained from the switchboard. The switchboard used was Panasonic's Smart Cosmo 21. The server used Ruby on Rails. The database used MySQL. The database holds not only the power data from the switchboard but also data of year, day and time. The server provides a scheduler application which runs on a web browser and visualizes the current consumption using the data from the database. The scheduler receives user operations on the Web browser and communicates with the server through the WebSocket connection. The server controls the home appliances by ECHONET Lite communication.

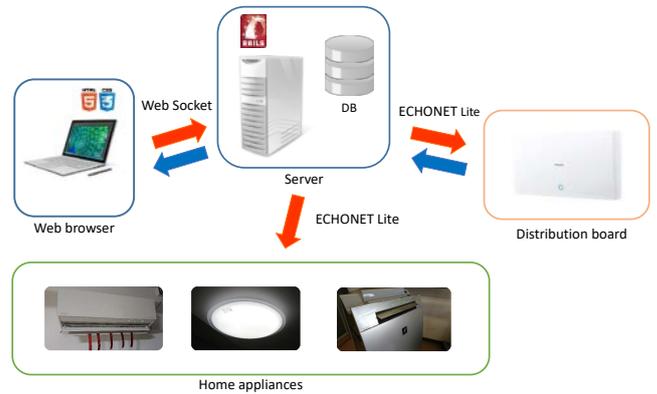


Figure 2: An entire system configuration.

3.1 Control flow

Figure 3 shows a sequence of visualization. Data of the electric current value used for the visualization are acquired from the switchboard every five minutes. The data of the current consumption are stored in the database. A user accesses the server by a Web browser. The browser requests the screen of the scheduler from the server. The server accesses the database and retrieves the data of the day which have been obtained until the current time as well as the data obtained in the previous day. The scheduler graphically displays the data on its main area.

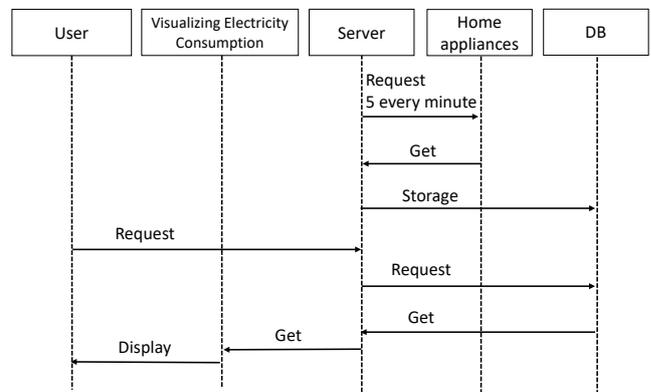


Figure 3: A sequence of visualization.

Figure 4 shows the control sequence. To use the controller, the user should initially drag an icon to the graph. When an icon is dropped onto the graph, the scheduler creates a sub area which shows a controller that matches the kind of icon and the position where the icon has been dropped. The controller requests the server to operate the home appliances. The server operates the home appliances using the ECHONET Lite protocol.

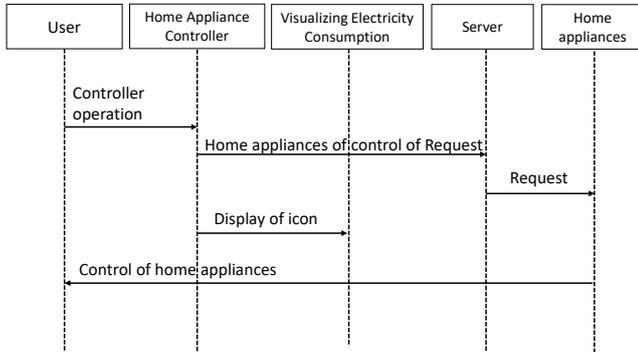


Figure 4: A sequence of control.

3.2 User interface

Figure 5 shows an example of the display screen. The scheduler functions as follows: Icons which represent home appliances are displayed, with each icon displayed in a color associated with the corresponding appliance. The data of the instantaneous current consumption of the day and those of the previous day are used in the graph. The consumption current of the day is indicated by yellow bars, while that of the previous day is shown by a blue line chart. A switching function for displaying only one of the two kinds of data is implemented. A vertical blue line is drawn at the current time to help users locate the latest consumption current at a glance among a lot of data displayed on the graph. The red broken line indicates the rated maximum current of the switchboard. Exceeding this maximum current results in a blackout causing all home appliances to stop operating. A likely reason for this situation is the use of too many home appliances or use of an appliance that consumes a high amount of current. By using the scheduler, users can prevent the blackout while promoting energy saving.

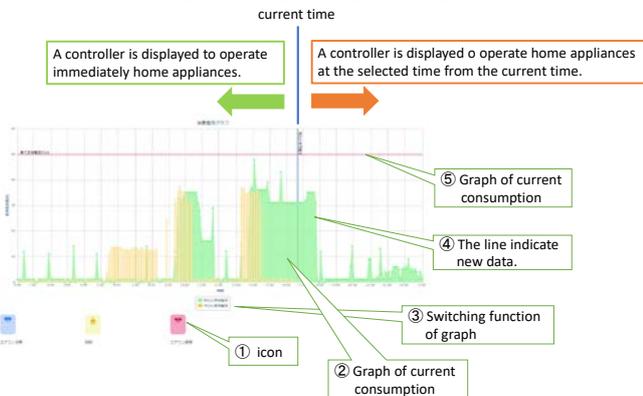


Figure 5: The functions of the main area.

3.3 Function of icons

Figure 6 shows the functions of the home appliance icons. The function to drop an icon onto the graph is implemented by using jQuery UI. The icons have the drag function, while the graph has the drop function. When an icon is dropped on the graph, the height of the icon changes according to the current consumption of the corresponding home appliance. The width of the icon can be changed by dragging its right edge. The icon has an initial width of one hour. As an additional function, increasing the width of the icon by one step lengthens the operation time of the corresponding home appliance by one hour. Double-clicking an icon in the graph deletes the icon. Dropping an icon on the left side of the vertical blue line indicates the current time invokes a controller for immediately operating the home appliances. Dropping an icon on the right side of the line invokes a controller for setting the power-on period of the appliances.

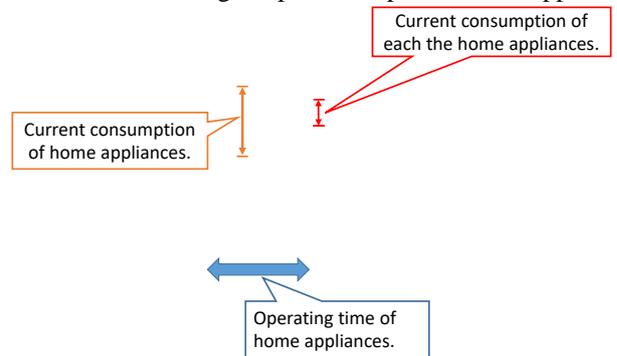


Figure 6: The functions of household appliance icon.

Figure 7 shows the home appliance controller in the sub area. Users can control a home appliance by clicking the control buttons. The controller sets basic operations of the buttons for the home appliance. As one example, we created a lighting controller and implemented the functions of turning on and off the lighting, controlling its illuminance, and discontinuing its operation after a specified time from the present time.

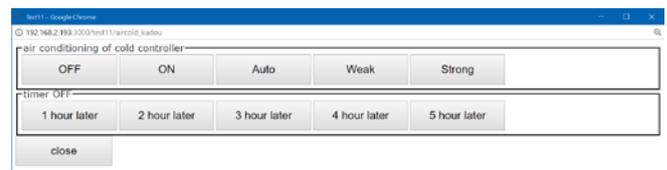


Figure 7: The control buttons in the sub area.

Figure 8 shows scheduling buttons in the sub area. We implemented a screen for programming a home appliance to begin operating after the selected time from the current time.

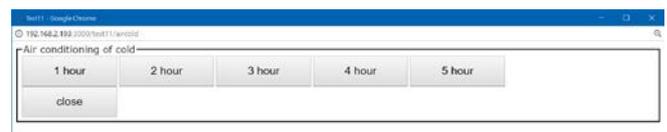


Figure 8: Scheduling buttons in the sub area.

4 EVALUATION OF THE SYSTEM

We have investigated the effectiveness of the scheduler. We conducted comparative experiments with the prior art and asked 23 students to join the experiments. We conducted a questionnaire survey on the interface. Two conventional types of interface, Toshiba Femininity Club and Panasonic AiSEG2, were used as the prior art [6][7].

Figure 9 shows the result of the questionnaire as to the home appliance operation using the icons and the visualization of the current consumption. Regarding the visualization of the home appliance operation using the icon width, the largest number of subjects said they were "satisfied". As for the visualization of the current consumption, the largest number of subjects said they were "very satisfied". These results confirm the effectiveness of the visualization using the shape change of the icon.

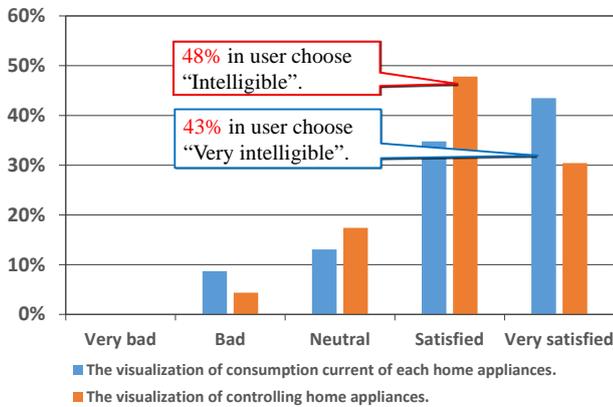


Figure 9: The result of visualization of the state of home appliances using icons.

Figure 10 shows the result of the questionnaire as to where to use the scheduler. The answers were counted for each of the two groups, i.e. users living with their parents and those living alone. The largest number of users living with their parents chose the "living room". As for the students living alone, the "school comminuting" and "living room" were the two most popular choices.

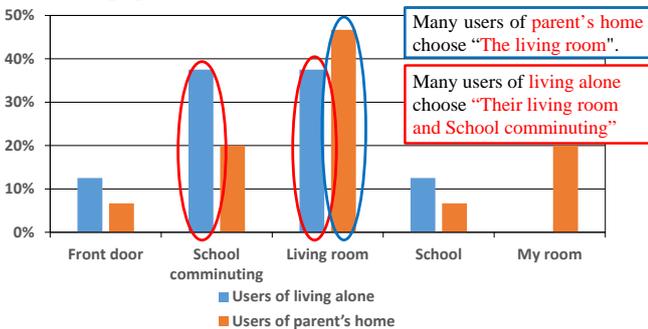


Figure 10: Users showed the place where they would like to use the scheduler.

Figure 11 shows the results of the questionnaire on continuous energy saving by using the scheduler. A majority of the users said that "continuous energy saving is possible". A higher percentage of the positive result was obtained from the users living alone.

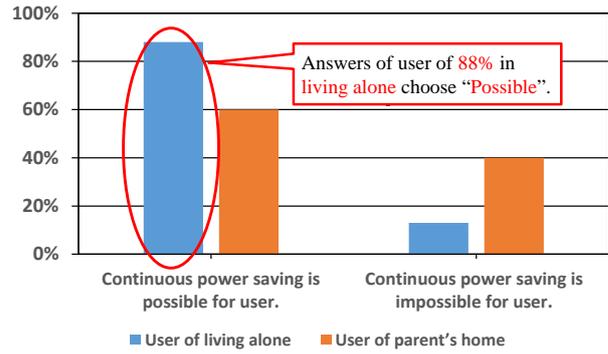


Figure 11: The results of the questionnaire on continuous energy saving by using the scheduler.

5 CONCLUSION

In this paper, we developed a scheduler to promote energy saving and evaluated its usefulness by experiments. The scheduler has the two functions of controlling home appliances and visualizing electricity consumption.

In future studies, we will quantitatively estimate how much the electricity consumption is reduced by the scheduler. We will also verify whether a continuous energy-saving effect can be obtained by the scheduler.

REFERENCES

- [1] Results of questionnaire of before and after use of "visualizing equipment" <http://warp.da.ndl.go.jp/info:ndljp/pid/10189884/www.city.osaka.lg.jp/kankyo/page/0000148884.html> (2016)
- [2] Investigation for mechanism to the improvement in added value of HEMS. http://www.env.go.jp/earth/house/conf/hems_04/mat02.pdf (2017)
- [3] Hidenori, Komatsu and Ken-ichiro, Nishio and Toshihiro, Mukai and Yasushi Shinohara, Automatic Generating System for Reports on Energy Conservation Tips Based on Electricity Demand Data: IEEJ Transactions on Electronics, Information and Systems, Vol.134 No.9 pp.1394-1405 (2014)
- [4] Takekazu, Kato and Kenji, Yuasa and Takashi Matsuyama, Energy On Demand : 2013 Information Processing Society of Japan, Vol.54 No.3 1185-1198 (2013)
- [5] Hiromu, Ohki and Morihiko, Tamai and Keiichi, Yasumoto, Interface for Supporting Energy Saving Plan to Achieve Energy Saving Goal without Reducing Comfort Level : IPSJ SIG Technical Report, Vol.2014-DPS-158 No.22 (2014)
- [6] Enrico C, Sarvapali R, Nicholas J, Understanding Domestic Energy Consumption through Interactive Visual-isation: a Field Study, Proc. of UbiComp 2012 (2012)
- [7] Toshiba Femininity club <http://femininity.toshiba.co.jp/femininity/>, (2017)
- [8] Panasonic Aiseg2 <http://www2.panasonic.biz/es/densetsu/aiseg/aiseg2/>, (2017)

Building Smart Block and Base Recognizing 3D Electric Circuit

Masaki Ohtani, Kazuki Hayashi, Kazuki Tokunaga, Junnosuke Takeuchi, Norikane Kanai, Masao Isshiki

Kanagawa Institute of Technology
ootani2014@ele.kanagawa-it.ac.jp

Abstract – IoT technology has been popularly applied in home environments, and home appliances connected to home networks are increasing. Home appliances connected to a home network are assigned respective IP addresses, and are managed by the home energy management system (HEMS). In addition, services for interactively operating home appliances connected to the network have been developed. In order to operate such household electrical appliances interactively, the position information of the installed home appliances is necessary. However, in the current home network system, it is not possible to recognize the installation position of home appliances. While Wi-Fi spreads, there is a technique for transmitting information signals through power lines (power line communication; PLC). We think that PLC communication can provide more information than Wi-Fi communication.

However, compared to Wi-Fi, PLC is less flexible in terms of the location of the appliance connection. To overcome this drawback, we have devised a “circuit block” system using blocks that can be horizontally and vertically arranged on a base. The base is equipped with a PLC interface, which provides 3D position information of each block placed on it. Each block has a built-in sensor and other components for the recognition of its vertical position. Combining these two kinds of information provides 3D position information of the block. We implemented a prototype of the circuit block system and evaluated the accuracy of the 3D position recognition. Possible applications of the circuit block system are also discussed.

Keywords: Block System, PLC communication, 3D position information, Interactive, Smart House

1. INTRODUCTION

The recent development of the IoT (Internet of Things) technology has enabled us to connect various products to communication networks. Particularly, the IoT technology has been popularly applied in home environments, and home appliances connected to home networks are increasing. Home appliances connected to a home network are assigned respective IP addresses, and are managed by the home energy management system (HEMS). It can be controlled via a tablet or mobile terminal by a communication protocol (ECHONET Lite) defined by the ECHONET Consortium [1]. In addition, services for interactively operating home appliances connected to the network have been developed. However, in order to operate such household electrical appliances interactively, the position information of the installed home appliances is necessary. Currently, the

position information of each home appliance is manually entered and fed to the HEMS.

While wireless communication by Wi-Fi has become widespread, there is power line communication (PLC) that transmits information signals through power lines. Previously, PLC was not popularly used in Japan, because the extremely high amount of noise in the power line prevented PLC from achieving high-speed data transmission. In recent years, these problems have been largely solved, and PLC-based commercial products have been put to practical use. With such a background, we have examined whether it is possible to physically obtain the location information of household electrical appliances by utilizing PLC's wired net.

Block systems have been widely used in engineering research. Many of them have microcomputers, sensors, and motors incorporated in the blocks. By combining these blocks, users can freely create various devices. Studies have also been reported on the recognition of the shape of a block-made structure and convert it to VR [2] [3] [4]. In these reports, the shape of the block-made structure is mainly recognized by two methods: a shape recognition method using a noncontact external sensor and a method of incorporating a sensor in each block and recognizing the positional relationship between the blocks.

In this paper, we propose a method to acquire 3D position information by combining PLC technology, sensors and blocks. For testing purposes, we made a system using a base and blocks designed for recognizing vertical and horizontal positions by two methods to obtain 3D position information. We report a system that obtains 3D position information by combining two methods. In addition, we discuss whether 3D position information by block type can be applied to smart house using PLC.

2. APPLICATION OF BLOCK SYSTEM

2.1 PLC communication for Smart House

PLC was originally developed for control of power network. Later on, its usage was expanded for general information transmission by mid-2000s. Figure 1 shows the conceptual diagram of PLC for residential use. As shown in the diagram, PLC allows a single power line installed in a house to be shared by both electric power supply and data signal transmission for an appliance connected to a power outlet. The biggest advantage is that complicated setup like Wi-Fi is not necessary [5]. However, despite the passage of more than ten years, PLC communication systems have not yet been widespread in Japan. The most serious problem is

the noise in the power line. In recent years, this problem has been largely solved and PLC communication systems will be more popularly used in the future. Meanwhile, future smart houses require a system for acquiring location information of home appliances. With such a technical background, we are studying whether it is possible to acquire location information circuitally by utilizing PLC. We will experimentally verify at the block levels.

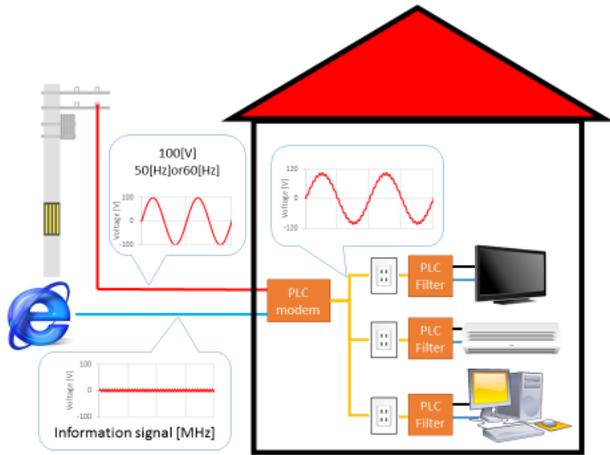


Figure 1. Principle of residential PLC.

2.2 Study of electronic block type tools

In Japan, products called "Electronic Block" are sold by Gakken [6]. Figure 2 shows one example of Electronic Block. Users can experimentally create a circuit by fitting blocks, in each of which conductors and electronic parts are embedded, into the base part.



Figure 2. The experiment electric block sold by Gakken, japan.

As a previous study, we have modeled a smart house energy management system utilizing a block system and created a power simulation device. Figure 3 shows the smart house power simulation system. We designed a "home appliance block" incorporating resistors in the block. Five types of home appliance blocks were designed for refrigerator, IH cooking heater, microwave oven, TV, and washing machine. In addition, each home appliance block was designed with two levels of power consumption: high and low (energy saving). A total of 10 home appliance blocks were produced. By controlling the current flowing

due to the difference in size of the resistors, power consumption was visually simulated. [7].

However, the fabricated device itself cannot recognize the connection position information of the installed sensors and motors. We think that there are two methods to recognize 3D position information of the sensors and motors. The first method is to recognize 3D position information from outside using a contactless sensor. The second method is to enable the sensors and motors to recognize their own 3D position information. Our idea is to combine these two methods. Furthermore, when this system is practically developed, it will be possible to acquire 3D position information of household electric appliances installed in a house.

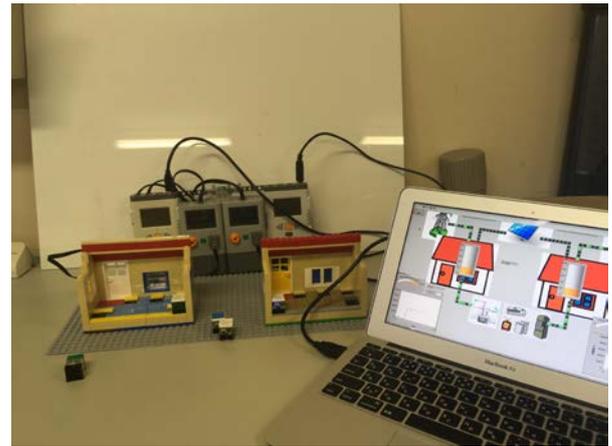


Figure 3. Electric power simulation system of smart house made with LEGO® block.

Most of such blocks offered so far are small blocks. Meanwhile, a large block was born in the United States. This block, called EVERBLOCK® (Figure 4), allows for creation of interior goods such as desks, chairs and simple huts. This block is a unique modular building system designed to bring out the designer in everyone [8]. If we incorporate electricity and network wiring into a big block like EVERBLOCK®, a more convenient intelligent block will be created.



Figure 4. Comparison between LEGO® block and EVERBLOCK®.

3. BUILDING OF 3D POSITION INFORMATION BY BLOCK SYSTEM

In this chapter, we will describe how to make a prototype smart base, wiring block, and electric circuit block. We considered a method for acquiring 3D position information of a block from two perspectives, i.e., the recognition of the vertical position and that of the horizontal position. For the development environment, we used National Instruments LabVIEW and NI myRIO. LabVIEW is the system development software most suitable for the development of systems for designing, prototyping, mounting and testing. NI myRIO is an onboard device with dual-core ARM® Cortex™ -A 9 real-time processor and customizable Xilinx FPGA I/O available.

Future smart houses will be wired from power and network communication lines installed under the floor or behind the wall. There is already a technology for this purpose: Power

Line Communications (PLC). PLC is a technology for superimposing data communication signals on domestic electric power (Figure 5) [5]. By introducing the PLC module, it is possible to connect home appliances to a network without using complicated wireless settings.

wiring technique that is intended to be applied in a management system for near future smart houses. Then we propose a block system to recognize position information at the block level. Figure 6 shows an image of the position information recognition block system. The “smart base” recognizes the two-dimensional position of each block. Further, a “wiring block” with a built-in wiring for power and data transmissions has been created. The communication between the smart base and electric circuit block uses the PLC technology. The block-level PLC is shown in the Figure 7. Similarly to Figure 5, both power and data are simultaneously transmitted through the same electrical wiring.

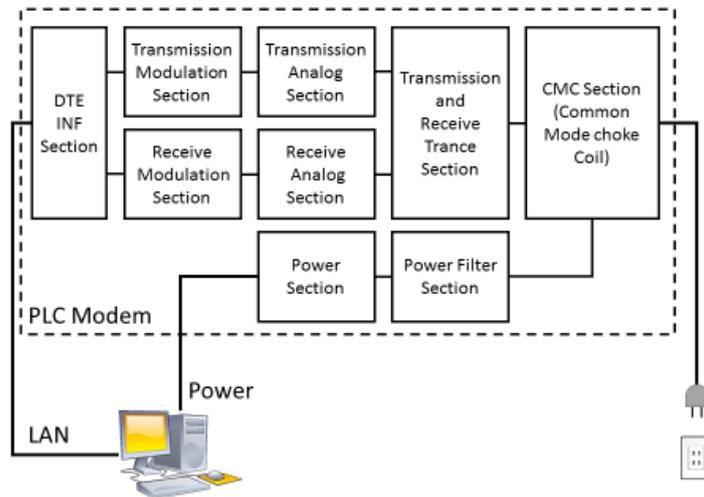


Figure 5. System configuration block diagram of PLC modem.

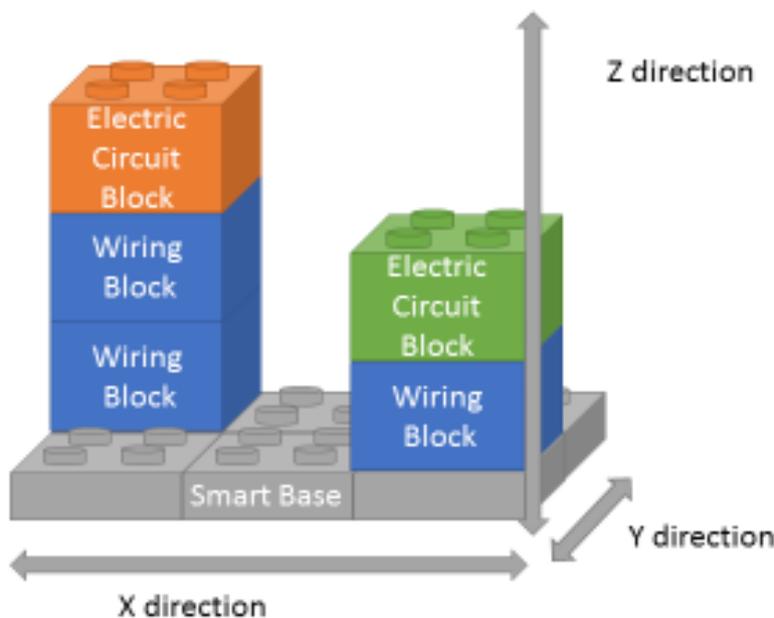


Figure 6. Image of block type position information acquisition.

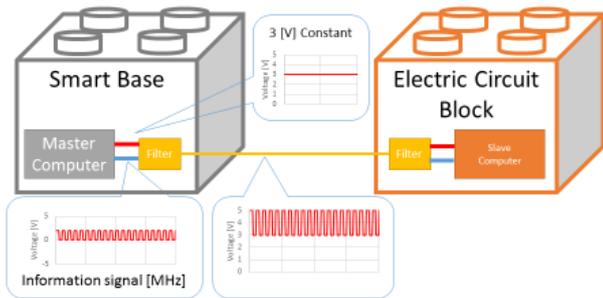


Figure 7. PLC of block level version.

3.1 Design of vertical position coordinate recognition circuit

For the recognition of the vertical position, we examined a method to represent the vertical position of a block by a potential difference given by parallel resistors. We tested a circuit as shown in Figure 8. In this circuit, the “electric circuit block” incorporates a microcomputer and two resistors-R1 and R2. The “wiring block” incorporating the resistor R3 allows the microcomputer in the topmost block to obtain information about the vertical level of this block by measuring a potential which changes with the number of wiring-block layers stacked on the base.

The value of the resistance was simulated under test conditions. Figure 9 shows the simulation results. The simulation revealed that setting the ratio of the resistance values R1: R2 to 1: 10 yields a sufficiently large potential difference for easy recognition of the vertical position. Furthermore, it was confirmed that the flowing current becomes small by using a large resistance. In the prototype, a block was made with R1 = 1 [kΩ] and R2 = 10 [kΩ].

3.2 Horizontal position coordinates and function recognition by image processin

We have developed a system that recognizes the horizontal coordinates as well as the function of each block using image processing. The prototype utilized the home

appliance blocks designed in 3.1. We assigned different colors to different types of household appliance blocks, acquired the state of the household electric appliance blocks with camera, and made a prototype system to recognize the position information and type of household appliance block using image processing. Figure 10 shows the prototype system. The household appliance blocks were placed on the black block serving as the base. The color of the household appliance block and its horizontal coordinates were recognized with the upper camera. In the color recognition method of image processing, each color was recognized by three elements of HSV (Hue, Saturation Chroma, Value Lightness Brightness). Figure 11 shows how the home electronics block coordinates are recognized. In the acquired image, the four colors of red, blue, yellow and green are respectively binarized and converted into a digital signal representing a matrix coordinate of 3 × 3. Furthermore, we built an electric circuit block incorporating a microcomputer and verified whether it can recognize the set coordinates and transmit power and data. For power transmission and communication method, the same operation as PLC was made possible by placing the signal between 3-5 [V] for a microcomputer driven between direct current (DC) 3-5 [V] by applying PLC technology.

3.3 Evaluation of 3D Position Information by Block System

The system produced in this report acquires three dimensional position information by separately recognizing vertical and horizontal positions. Therefore, we evaluate the two systems.

The vertical coordinate recognition circuit requires actually embedding the circuits into blocks. What problems will occur when those blocks are actually formed remains to be studied.

Horizontal coordinate recognition by image processing has the problem that two adjacent blocks of the same color cannot be identified as two blocks.

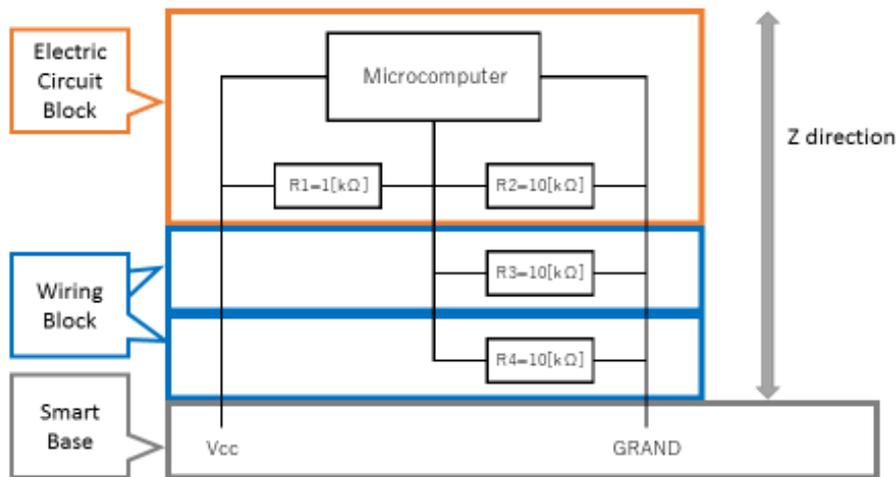


Figure 8. Design of vertical coordinate recognition circuit.

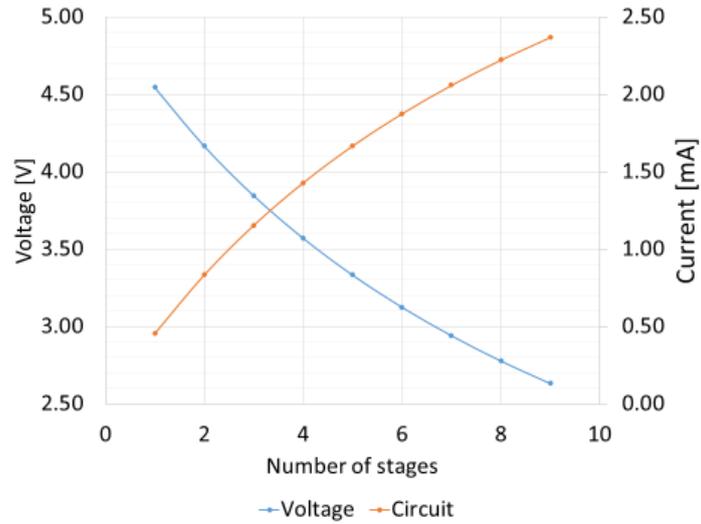


Figure 9. Result of simulation of vertical coordinate recognition circuit.

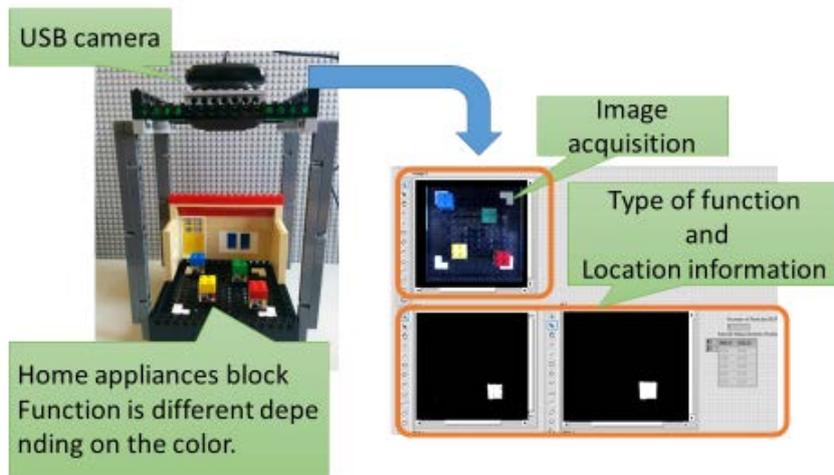


Figure 10. Recognition of position information of household appliance block by image processing.

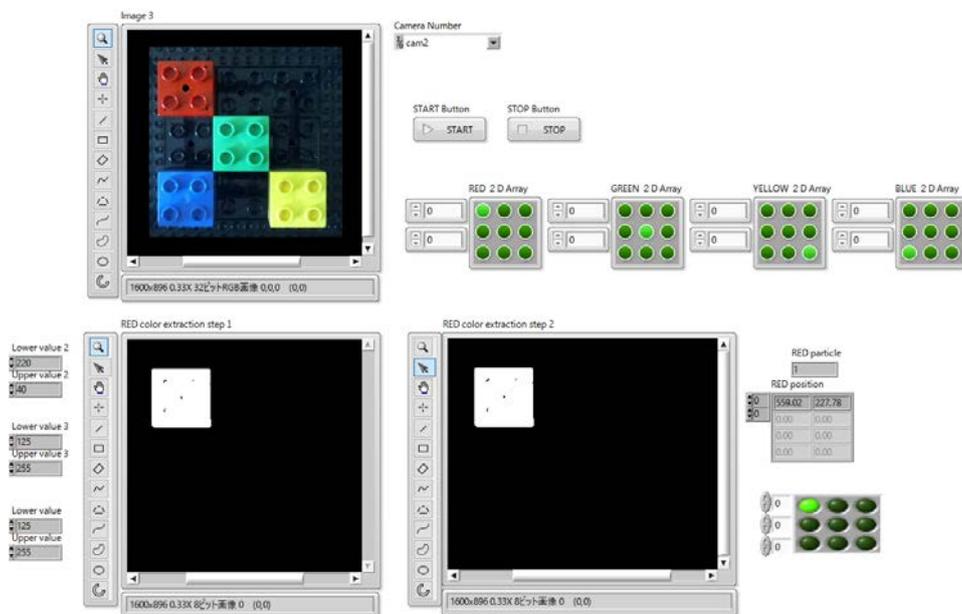


Figure 11. Convert position information to digital signal.

4. DISCUSSION FOR NEXT STEP

In this chapter, we will discuss the next step of the smart block that recognizes 3D position information.

USB interface can recognize the type of connected device. Currently, USB 3.1 is available, along with the Type-C connector which has advanced miniaturization of the connecting terminals. Type-C is equipped with a Power Delivery controller IC, which enables high power reception of 100 W. Furthermore, high-speed data communication by the Biphase Marked Coding (BMC) method has been available. As a result, computers, monitors, printers, and other peripheral equipment that previously required separate power supplies can now be powered by a single AC adapter. In the future, as the receivable amount of electricity further increases, more home electric appliances will be powered through USB connections and the household outlet will be changed to USB interface. Future houses will probably be “smart” enough to recognize the type and function of home electric appliances in addition to supplying them with power upon connection of such appliances to the outlets. We believe that our block systems that recognize 3D position information can be further advanced by applying USB power delivery and PLC technology. Furthermore, we would like to develop an “intelligent wall” that recognizes position information of household electric appliances by utilizing a large block such as EVERBLOCK®. Currently, we are designing intelligent walls using EVERBLOCK®. An intelligent wall under test is shown in Figure 12. We incorporated a power supply, microcomputer and LED into the EVERBLOCK®.

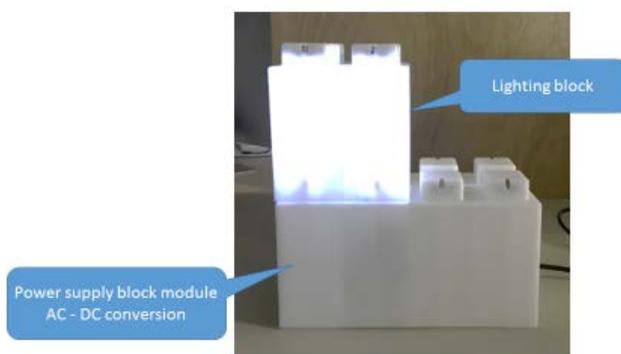


Figure 12. Prototype of intelligent wall.

5. CONCLUSION

In this report, we proposed a system to recognize the position information of blocks. In order to recognize the coordinates of each block, we created systems for horizontal and vertical position recognitions. For the horizontal position recognition, block colors were recognized using the image processing, and 2D coordinates could be obtained. For the vertical position recognition, the potential difference produced by parallel resistors was utilized. The two kinds of position information were combined to obtain 3D position information. In the future, we would like to experimentally explore the possibility of interactively operating multiple

circuit blocks that can acquire the 3D coordinate information.

REFERENCES

- [1] ECHONET consortium web page, <http://echonet.jp/>, (2017.5).
- [2] J. Nagai, T. Numano, T. Higashi, M. Tessier, K. Miyata, An interactive VR system using toy blocks and its evaluation, Entertainment Computing Symposium, (2013).
- [3] M. Ando, T. Hosoi, K. Nakajima, K. Takashima, Y. Itoh, T. Adachi, A. Onoye, Y. Kitamura, StackBlock: Building Block Interface Recognizing 3D Shape of Stacked Blocks, Information Processing Society of Japan Interaction,(2014).
- [4] Y. Itoh, Y. Kitamura, M. Kawai, F. Kishino, ActiveCube: A Bi-directional Interface for Real-time Constructing and Interacting with a 3D Environment, Information Processing society of Japan Journal, (2001).
- [5] Institute of Electrical Engineers of Japan, High-speed power line communication system and EMC investigation expert committee, High-speed power line communication system (PLC) and EMC, Ohmsha Publisher of science and Engineering books, (2007.11.20).
- [6] GAKKEN HOLDINGS CO.,LTD. web page, Gakken Electronic Block EX , <http://www.gakken.co.jp/>, (2017.5).
- [7] R. Uda, Y. Ymada, M. Ohtani, Y. Ikeda and N. Kanai, Development and application of Block Type Teaching Materials for Introductions of Engineering, Japan Society for Engineering Education, (2017).
- [8] EVERBLOCK® web page, <http://www.everblocksystems.com/>, (2017.5).

Acceleration Method for Real-time Processing in Distributed Control System

Akihiro Yamashita^{*}, Hiroshi Mineno^{**}, and Tadanori Mizuno^{***}

^{*}Graduate School of Informatics, Shizuoka University, Japan

^{**}Faculty of Informatics, Shizuoka University, Japan

^{***}Faculty of Information Science, Aichi Institute of Technology, Japan

Yamashita.Akihiro@ma.mee.co.jp, mineno@inf.shizuoka.ac.jp, mizuno@mizulab.net

Abstract – Overall performance enhancement is pursued for the distributed real-time control system by improving performance and using sensor information for control operation. However, to achieve high performance for the real-time control system, issues such as interrupts occurring frequently due to the shortening of the control cycle, and the access to and from slow devices preventing advanced microprocessors from working at full potential have not been addressed. In this paper, we propose approaches to accelerate processing of data by layering of the system bus architecture and the memories, and transferring I/O data by DMA transfer. We achieved 1.6 times higher processing capacity for the control system by improving the hardware architecture based on the analysis and evaluation of system behavior using the information on the hardware configuration and the actual data logs.

Keywords: microprocessor, bus architecture, real-time processing, distributed control system, embedded controller.

1 INTRODUCTION

The acceleration of real-time processing in distributed control systems was initially driven by the increased performance of microprocessors and high-speed memory access. In the meantime, the clock operating frequency has been increased by the change in the architecture of microprocessors from the complex instruction set computer (CISC) to the reduced instruction set computer (RISC) [1], [2]. Along with the acceleration of the microprocessor processing speed, different types of access methods have been recommended for the dynamic random access memory (DRAM) to accelerate the transfer of instruction codes to a microprocessor. Block transfer operations move the execution programs stored in a DRAM to the pipelines via the instruction cache memory of the microprocessor. The cache memory operates at the same frequency as the microprocessor does, which is the factor that contributed to the acceleration of microprocessors. Moreover, storing the executed programs in the cache memory enables high-speed processing because instructions are executed every clock cycle when the same programs are executed.

In a real-time processing system with an embedded controller, conditional branching or interrupts are frequently used in programming. When a cache miss occurs, effectiveness of the instruction cache memory is reduced and the potential of the microprocessor is not fully exploited. To reduce such cache misses, a level 2 (L2) cache memory with a larger capacity was adopted to achieve higher performance of embedded controllers. However, it has become difficult to further increase the performance only by increasing the clock speed and cache memories. It is now

critical to pursue higher performance by further enhancing the system architecture of embedded controllers. In this paper, we report on the limits of evolution of controllers only by improving microprocessors, propose an architecture to achieve higher performance, and report on the implementation and evaluation of the proposed architecture.

2 CONVENTIONAL TECHNOLOGIES AND ISSUES

2.1 Increase of the Clock Frequency and the Use of Cache Memories in the Microprocessor

Since the 1990s, the microprocessor architecture to simplify the instruction set execution has been studied for the faster clock speed. The RISC processors have been recommended to improve the operation speed by simplifying the architecture with reduced instruction sets. In the RISC architecture, a shorter instruction decoding time is required and the pipeline efficiency is improved by the instruction read-ahead technique. As a result, each operation is executed in one clock cycle. In the microprocessor, an instruction is read and decoded (interpreted), data are taken out of the memory as required for execution of the instruction, and the result is stored in the specified location. The read-ahead, decoding, and execution operations in the pipelines are carried out in parallel, and each instruction is executed in one cycle. In the RISC processor, only the load instructions and store instructions affect the memory. Furthermore, the internal registers are increased as compared to those in the conventional processors. By storing the data being processed in the registers, the memory accesses are reduced.

Since the memory access speed has not so much increased as the microprocessor operation speed, cache memories with the operation speed as fast as the microprocessor clock speed were integrated in the microprocessor. There are two types of cache memories: the instruction cache memory and the data cache memory. The instruction cache memory stores instructions read by the read-ahead operation to the pipelines. When the same instructions are read, they are read from the cache memory. Since the cache memory in the microprocessor has a limited capacity, older instruction codes are sequentially replaced with newly read codes. By storing the programs (instruction codes) that are repeatedly executed in the cache memory, instructions are efficiently transferred to the pipelines. The full potential of the pipelines is exploited, achieving the acceleration of the

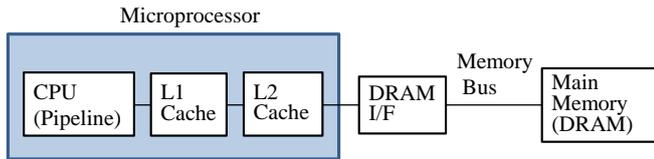


Figure 1: Microprocessor and Memory Configuration

microprocessor. Furthermore, the data cache memory is added to store the data once read from the external memory, which are required for the operation. As is the case with the instruction cache memory, by storing the data that are repeatedly used in the data cache memory, the data reading time is shortened and the microprocessor operations are accelerated. When the next instruction is not found in the instruction cache memory, a cache miss occurs and the instruction is read from the external memory. The cache miss time is the sum of the fetch time for the first data of the memory data block and the transfer time of the remaining data. To enhance the system performance of the RISC processor, it is necessary to increase the cache memory capacity and reduce cache misses. Due to cost limitations for the capacity of the high-speed cache memory in the microprocessor, the level 2 (L2) cache memory is adopted. In the microprocessor with an internal L2 cache memory, which was developed in the 2000s, the operation speed is further accelerated [3].

Figure 1 shows the configuration of the microprocessor with the internal level 1 (L1) and level 2 (L2) cache memories, and the external main memory. The use of an L2 cache memory reduces the reloading time loss due to cache misses of the L1 cache memory, and increases the throughput of microprocessor that operates at high drive frequencies. Furthermore, by the high-speed block data transfer from the external main memory to the L2 cache memory, the time loss of the CPU pipelines is shortened.

2.2 Acceleration of the Transfer Rate of the Dynamic Random Access Memory (DRAM)

For the DRAM external memory, the burst transfer method was developed. The address data lines are used both for the row address and the column address to specify the memory cell of DRAM. The row address and the column address are set in a time sharing operation. The upper bits of the memory address are used for the row address, and the lower bits are used for the column address. In the extended data output DRAM (EDO DRAM), data latches are provided for the data output lines to accelerate transfer by overlapping timings of the data output and the next row address reception. With the synchronous dynamic random access memory (SDRAM), developed after the above asynchronous DRAM interfaces, high-speed burst transfer is enabled by the synchronous operation with the system bus. With the double-data-rate SDRAM (DDR SDRAM), developed in the 2000s as another type of SDRAM, the data transfer rate is doubled relative to the standard SDRAM by transferring data on both the rising and falling edges of the clock signal. With DDR2 SDRAM, the

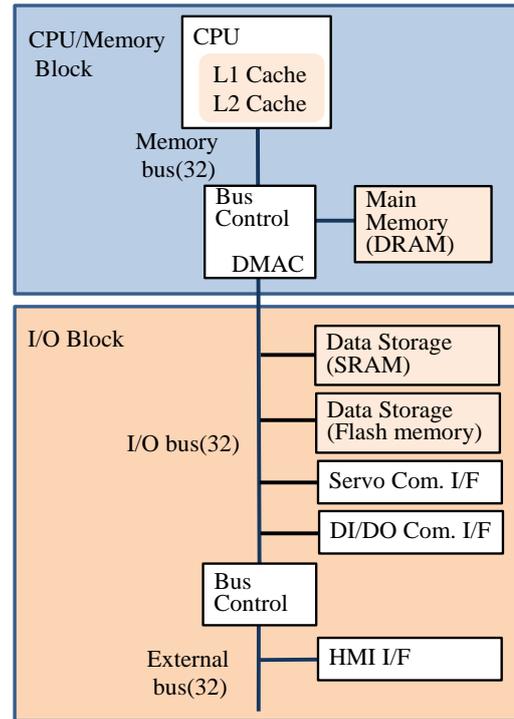


Figure 2: Block Configuration of the Conventional Embedded Controller.

data transfer rate is quadrupled relative to the standard SDRAM by doubling the external synchronous clock rate of DDR SDRAM. Furthermore, with the double-data-rate3 SDRAM (DDR3 SDRAM), the data transfer rate is octupled relative to the standard SDRAM by quadrupling the external synchronous clock rate of DDR SDRAM.

2.3 Configuration of the Real-time Processing System

Figure 2 shows the block configuration of the embedded control system that performs servo control as an example of real-time processing system. The embedded controller consists of the CPU/memory block and the I/O block. The CPU/memory block consists of the following components: the microprocessor with the internal L1 and L2 cache memories, main memory (DRAM), and bus control section. The CPU is connected to the main memory (DRAM) with a high-speed memory bus via the DRAM interface.

In the I/O block, the non-volatile memory (flash ROM) storing system programs and the SRAM storing user programs and parameters are connected to the I/O bus. In order to control the external devices, the servo communication interface and the DI/DO communication interface are also connected to the I/O bus. The I/O bus is connected to the bus control section, from which it is connected to the external devices as the external bus. The human-machine interface (HMI) is connected to the external bus. The HMI interface is used to output the information or status controlled by the controller to the display, and to input the switch information of the operation panel for machine control. In the distributed control system for servo control, multi-CPU control is performed for command generation,

servo positioning control, machine I/O control, display control, and communication control [4].

The following section describes system operations. At power-on, system programs are read from the non-volatile flash memory, and stored in the main memory. The system programs in the main memory are then transferred in block to the L2 cache memory, L1 cache memory, and pipelines of the CPU where programs are executed. The controller internal states or the control data to be retained are stored in SRAM.

The data received through the servo communication interface are transferred to the main memory by DMA transfer using the DMA function of the bus control section of the CPU/memory block. The command data for servo control are transferred from the main memory to the servo communication interface by DMA transfer. The status information of external devices are read by the CPU through the DI/DO communication interface. The control signal is output from the CPU to the external devices through the DI/DO communication interface. These DI/DO control data are processed by the program I/O function of the CPU.

2.4 Characteristics of the Embedded Control System

2.4.1 Interrupt processing

In the embedded control system, real-time control is performed by reading the condition of external devices using sensors and controlling the actuator using programs. Real-time control is used to generate interrupts at a fixed cycle, switch programs according to the interrupt cycle, read the external conditions, process data to give commands to the controlled objects, and generate physical I/O signals. In order to control the objects, this series of operations must be completed within a certain time limit.

Acceleration of the series of operations, from reading of the external conditions, command data processing, to control command signal output, is required in order to improve the control performance of this embedded control system. Furthermore, to satisfy the needs for more advanced control functions, the volume of information of the external input must be increased, and more control output signals must be sent to the controlled objects. In order to enhance functions and accelerate operations, the control system must have high-speed data processing performance to allow processing of the increasing amount of I/O information in a short time.

Figure 3 shows an example of servo control with an arc-shaped path using the embedded control system. To perform position control for the movement from A to B2 at the speed F in the unit time Δt , the movement amount for each of the X and Y axes is the product of the speed component F_x/F_y and the time Δt . When Δt is reduced by one half while the speed F remains the same, the command is given for the movement from A to B1, and then for the movement from B1 to B2. Therefore, the path control accuracy is increased, enabling higher servo control accuracy.

The response of servo control can be improved by shortening the output cycle of the control command.

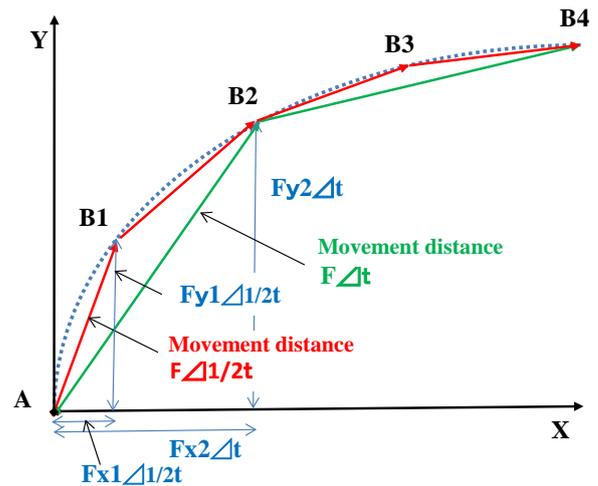


Figure 3: Movement Command and Path Control for Servo Control.

Alternatively, higher control is enabled by increasing the number of control axes while the response remains the same. The fixed-cycle processing is performed by switching the control software operation using interrupts in synchronization with the cycle.

In the meantime, programs executed by the microprocessor are suspended due to the interrupt control. Cache misses occur in the cache memory that transfers commands to the pipelines of the microprocessor. Therefore, commands must be transferred from the external DRAM to the pipelines and the cache memory. In the controller for the real-time control, every time tasks are switched by interrupt, handling of the program's cache misses is required.

Figure 4 shows an example of analysis of the control cycle and the breakdown of the CPU processing time. The vertical axis shows the generation of the control system, and the horizontal axis shows the interpolation period for the control operation.

The CPU portion of the processing time for the third generation system corresponds to the time period required for the microprocessor to perform the cache memory program for data processing. The DRAM portion corresponds to the time period required for the microprocessor to read programs from the main memory. The I/O portion corresponds to the program I/O processing time.

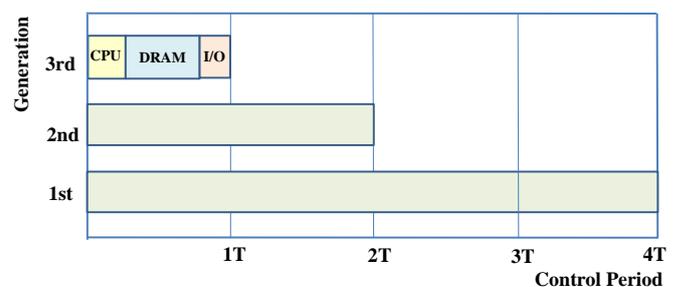


Figure 4: Example of the Analysis of Control Cycle and CPU Processing Time.

The interpolation period for the control operation is reduced by one half for each generation change, which shows enhancement of the processing performance of control system. Since the CISC processor of the first generation system was changed to the RISC processor for the second generation system, the processing performance was enhanced by the increase of drive frequency and the use of cache memory. For the third generation system, the microprocessor with L2 cache memory was adopted. The further increase of the CPU drive frequency enabled enhancement of the processing performance.

2.4.2 Control I/O processing

The embedded control system detects real-time changes of the external devices by input of external information, performs necessary control operations, and outputs control signals to the external devices. For real-time control, input of information and output of control signals from and to the external devices are performed in a fixed cycle. To enhance the functions, the volume of input information to be processed and the volume of control output information are increased [5].

The embedded control system processes data to generate control commands, and performs I/O control for input of external conditions, output of control signals, etc. Those functions are divided according to the processing time. The control system accesses external sensors or actuators, which are quite slow devices when compared to the microprocessor with cache memories for data processing. In order to prevent degradation of the microprocessor performance, it is necessary to reduce direct access by the microprocessor program I/O to and from the slow devices. For that purpose, data in the I/O data buffer are transferred to the external devices using serial communication [6], [7], [8]. In this method, I/O data are stored in the external memory, a DRAM, to allow relatively fast access, which enables I/O control of the data access to and from slow devices using serial communication. Direct access can be avoided by using this method, and degradation of the system performance is mitigated.

2.5 Issues to Accelerate the Distributed Real-time System

As mentioned above, performance of the embedded real-time control system was enhanced by adopting the technologies to achieve increases of the microprocessor's operating frequency, larger capacity of the internal cache memory, and acceleration of the block transfer rate of the external DRAM memory until the 2000s. For I/O data processing, the load on the main CPU was lightened by using serial communication to distribute control for servo control or DI/DO control. However, the following issues must be addressed to achieve further enhancement of the performance.

Issue 1

Performance and functions of the real-time control system can be enhanced by shortening the control cycle. However, the shorter control cycle increases the number of interrupts

per unit time, and the task switching increases rereading of programs from the external memory [9]. Based on the analysis of data processing in the third generation system shown in Figure 4, rereading of programs from DRAM accounts for significant part of the CPU processing time. The issue is that the control performance of the controller cannot be improved by accelerating the microprocessor's operating frequency.

Issue 2

The real-time control system requires frequent I/O data access to and from external devices, as well as those in the above-mentioned main memory. As mentioned above, the load on the microprocessor is lightened by using serial communication to distribute control for input of external condition signals or read data and output of control signals or data to external devices. However, the processing speed for the program I/O to and from the external devices is significantly slow when compared to the processing speed in the cache memory. The issue is that the throughput of microprocessor is lowered by the slow processing speed for the program I/O via the communication buffer or memory.

3 PROPOSED METHOD

3.1 Restrictions in the Distributed Real-time System

To create a system that enhances performance for real-time control, it is necessary to consider trade-off between acceleration and cost/power efficiency. Performance of the controller has been enhanced by the increase of L1 cache memory capacity and the use of L2 cache memory in the microprocessor, and acceleration of block transfer rate of DRAM. However, due to time loss for rereading after cache misses, it is difficult to enhance system performance only by accelerating the clock cycle of the microprocessor.

Although this problem can be solved by reducing the program size for control tasks that are repeatedly used so that they can be stored in the cache memory, a heavy burden will be placed on the software designers. The cache memory capacity may be increased to solve the problem, but the cost will be increased. For cost/power efficiency, assuming the use of a high-performance commercial microprocessor and memory (DRAM), hardware architecture is a key factor in obtaining utmost performance from each of the system elements. In this paper, in order to enhance the system performance, we propose layering of the memories and the bus architecture to suit the characteristics of the controller, focusing not on the control software structure, but on the hardware.

3.2 Hardware Architecture

Figure 5 shows the configuration of distributed control system we propose in this paper. The control system for servo control has three types of memories: the flash ROM memory to store system programs, SRAM memory to store user programs and parameter data, and main memory (SDRAM). Data are input or output via the servo

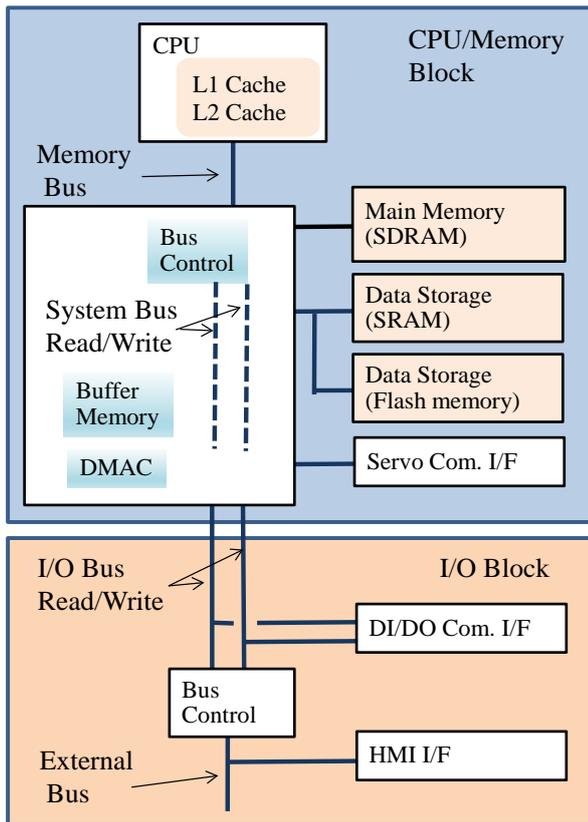


Figure 5: Block Configuration of the Proposed Distributed Control System.

communication interface for communication with servo amplifiers for servo control, via the DI/DO communication interface for machine I/O control, via the HMI interface for communication with a terminal for machine operation and status display. The system has four types of buses: the memory bus, system bus, I/O bus, and external bus.

3.3 Access Time Reduction by Layering Memories

In order to optimize the memory access, the memories in the system are layered according to the difference of the access speed. When programs are read from the microprocessor or the main memory, high-speed block transfer is performed between the cache memory and the main memory. Increase of the efficiency of block transfer requires optimization of the data size.

In order to ensure reliability of the real-time control, the write through cache method is adopted for writing data to the cache memory. When the write through cache method is used, all data must be transferred to the memory or I/O device in the layer right below. When this data transfer is completed, the CPU determines that writing is done. In this case, the CPU waits until writing to the main memory, SRAM, or I/O device is completed. Since SRAM or I/O devices are slow devices as compared to the main memory, access time becomes longer. In this paper, we propose the use of a high-speed buffer memory, which makes the write through cache method feasible. When the buffer memory is

provided, the time required for writing can be shortened since the CPU can proceed to the next process after writing data to the buffer memory, without being affected by the access to the slow devices. The contents of the buffer memory are transferred to the SRAM or I/O device by DMA transfer.

3.4 Layered Bus Architecture and Parallel Execution of the DMA Transfer

In order to enable high-speed data transfer from the main memory to the microprocessor, a dedicated memory bus is used for the connection between the main memory and the microprocessor. For the connection between the main memory and other memories or communication interfaces, the system bus is provided. Furthermore, the I/O bus is provided for I/O control, and the external bus is provided for the connection with external devices. The bus architecture is configured in the system using the layers of four types of bus according to the difference of access speed and the bus width.

The memory bus used for the main memory is separated from other data buses used for I/O control. DMA transfer must be enabled for data I/O between the main memory and the SRAM or communication interface, which are connected to the system bus or I/O bus. Thus, while no access is required to the system bus or I/O bus by the CPU, the control data I/O operation can be performed simultaneously with the CPU operation. Furthermore, the layered bus architecture enables separation between the I/O devices that require fast data transfer for large volume data and the I/O devices with non-frequent access and low throughput requirements. Optimization of access to the memories or I/O interfaces can be achieved by layering of data buses.

For the system bus and the I/O data bus, there are two types of buses, read only and write only. Contention with the main CPU on the data bus can be avoided by having different buses for CPU reading access and DMA transfer access. This approach for bus separation is also effective for the control system consisting of multiple CPUs or the system in which the bus is occupied by DMA transfer.

4 IMPLEMENTATION SCHEME

4.1 Software Architecture

Figure 6 shows the configuration of software tasks in the embedded controller. The system software consists of the basic real-time operating system and the programs running on it to perform the following tasks: synchronization control of the system; pre-processing analysis of the commands in the processing programs to apply the result to the position control command and the speed command; motion control to send the positioning servo command for machine operations; machine I/O control to execute sequence programs for machine operation control, read status signals from external devices, and output commands required for machine control; operation control to read the switch information for machine operation and output the machine

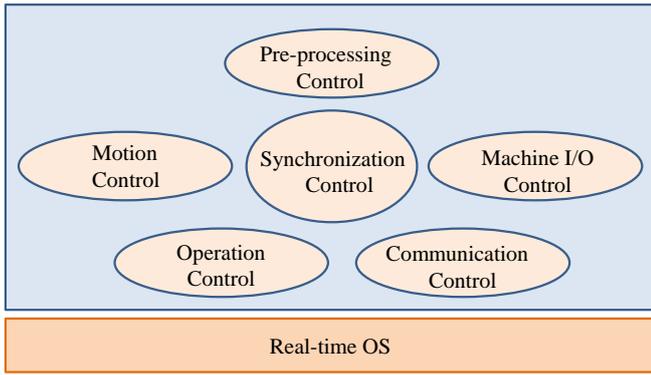


Figure 6: Software Configuration of the Control System.

movement of controller status information to the display; and communication control for information exchange with the external devices.

For real-time control, these tasks are switched and performed in the fixed control cycle. In addition, interrupts are issued as appropriate when a number of events occur. When task switching or an interrupt occurs, the flow of system programs is changed by context switching, and cache misses occur in the cache memory. Cache misses also occur when the program size is large relative to the size of the L1/L2 cache memory. The above-mentioned factors cause cache misses. When a cache miss occurs, system programs are read from the main memory and the CPU is stopped while this occurs.

4.2 Efficiency Improvement of the Cache Memories

Figure 7 shows the configuration of the CPU, L1 cache memory, and L2 cache memory in the microprocessor and the main memory. The main memory is high-speed DDR-DRAM. Programs stored in the main memory are transferred to the pipelines of the CPU. The L1 cache memory operates at the same frequency as the CPU clock frequency. The L2 cache memory generally operates at half the speed of the L1 cache memory. High-speed block transfer is performed between the main memory and the L2 cache memory. The block transfer rate is determined by the performance of the memory (DDR-DRAM). The transfer data size is variable. When the data size is increased, the data transfer efficiency is improved, but it takes more time to transfer data. It is necessary to consider the trade-off between the system program block size and the cache miss occurrence frequency. For writing, a high-speed buffer memory is provided to use the write through cache method. The cache memory efficiency is improved and the CPU write cycle is shortened by using the buffer memory. The effective use of the buffer memory will be described later.

4.3 Layering of the Memories

Figure 8 shows the configuration of the CPU/memory block with the memory layers and the system bus architecture in the embedded control system. The system has the L1 cache memory and L2 cache memory in the

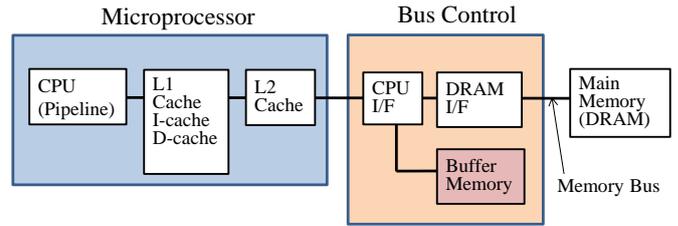


Figure 7: CPU and Cache Memory Layers.

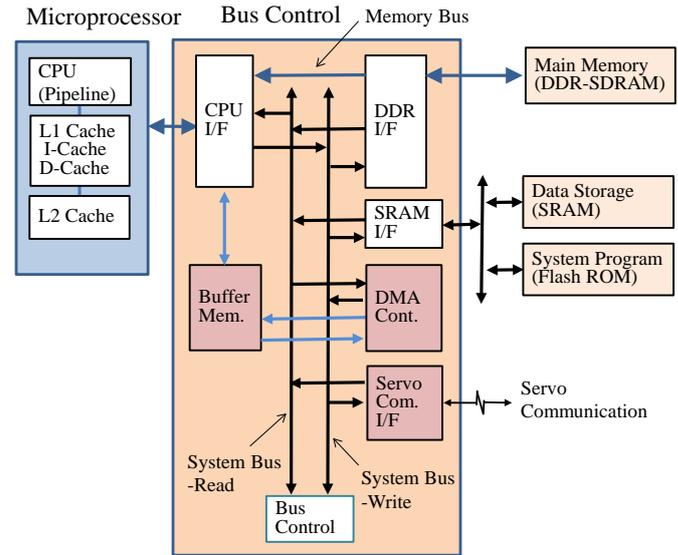


Figure 8: Memory Layers and System Buses in the CPU/Memory Block.

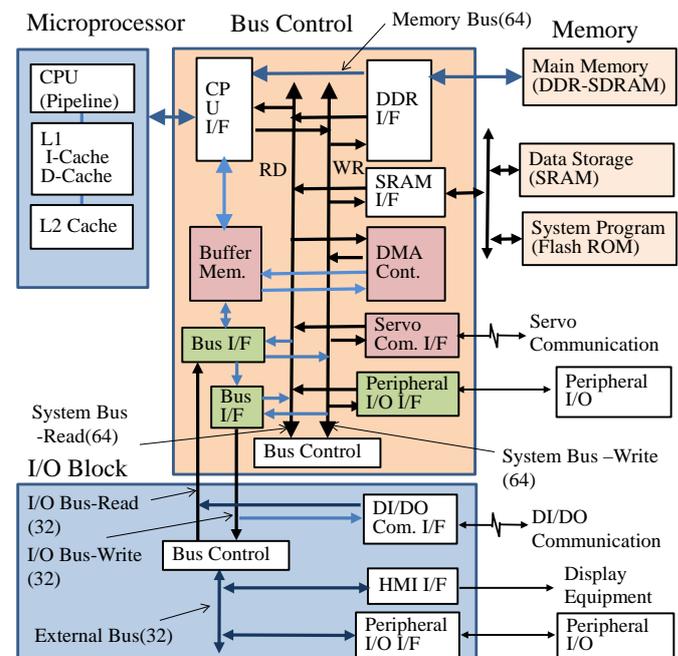


Figure 9: Bus Architecture of the Control System.

microprocessor, buffer memory, main memory (DDR-DRAM), flash ROM storing system programs, and SRAM storing user programs and parameters.

The cache memory is connected to the main memory (DDR-DRAM) via the CPU interface and the DDR interface using the dedicated 64-bit-wide memory bus. The data cache memory is connected to the buffer memory via the CPU interface using the 64-bit-wide bus. The SRAM and flash ROM are connected to the 32-bit-wide system bus via the SRAM interface. Data transfer between the buffer memory and SRAM is performed by DMA transfer. These memories are layered according to the data read/write access speed and the bus width.

Then the layered memories are connected to the layered buses to optimize the system performance.

4.4 Layering of the Data Bus Architecture

Figure 9 shows bus layers of the control system. Data buses are layered in the control system. In addition to the above-mentioned memory bus and system bus, the system has the I/O bus for the control device connection, and the external bus for the connection with the external devices. The buffer memory, DMA controller, servo communication interface, peripheral I/O interface, and I/O bus interface are connected to the system bus. The I/O bus is used to connect the DI/DO communication interface for machine I/O control. Then, the I/O bus is connected to the bus control block, from which it is connected to the external devices as the standard external bus. A standard HMI interface and standard peripheral I/O device are connected to the external bus. The DMA controller located on the system bus enables data transfer independent from the CPU.

5 EVALUATION

5.1 Method and Conditions of Evaluation

We conducted performance analysis and evaluation for the hardware architecture proposed in Section 4 in the following procedure.

- 1) Selection of the programs to be subjected to the evaluation
- 2) Collection and analysis of the system behavior logs
- 3) System performance check

The evaluation conditions are as follows.

- 1) Application of layering of the memories and buses as proposed in Section 4
- 2) Use of the same clock operating frequency for the microprocessor
- 3) Use of the same capacity for the L1 and L2 cache memories in the microprocessor
- 4) Use of the same data transfer width for the main memory. Acceleration of the transfer rate from 100 MHz to 133 MHz.
- 5) Acceleration of the clock operating frequency of the system bus from 100 MHz to 133 MHz
- 6) Creation of the log information collection program to be used for the system performance evaluation using the tasks for basic control: synchronization control, pre-

processing control, motion control, and machine I/O Control, which are selected from the control tasks shown in Figure 6

5.2 Evaluation Result

5.2.1 Collection and analysis of the system behavior logs

Bus transactions are measured using a logic analyzer, bus analysis tool, and aggregation tool. We measured the average number of accesses to each hardware resource in a fixed cycle for the selected program. Figure 10 shows the number of times accessed.

We then calculated the latency, which is the time required for access to the hardware resource, using the hardware configuration information, log information collection program, and the aggregation tool. Based on the calculation result and the times accessed, we made analysis on the time consumed by the hardware resources. Figure 11 shows the result of analysis of the consumed time.

Since cache misses occurred frequently, access for reading from the main memory (SDRAM) accounts for 60%. We also found that much time is consumed for reading from SRAM due to large latency despite the low number of accesses. Based on the above-mentioned result of measurement and analysis of the consumed time,

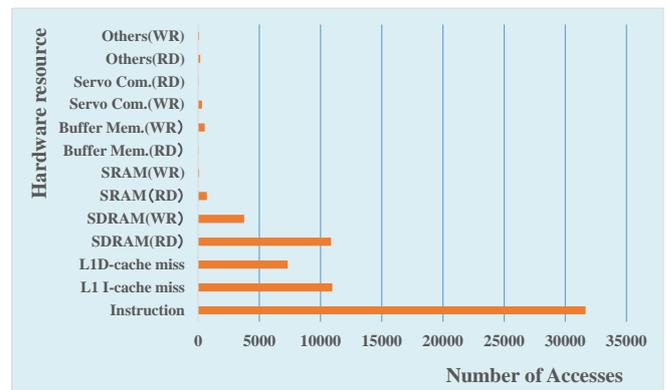


Figure 10: Number of Accesses by Hardware Resource.

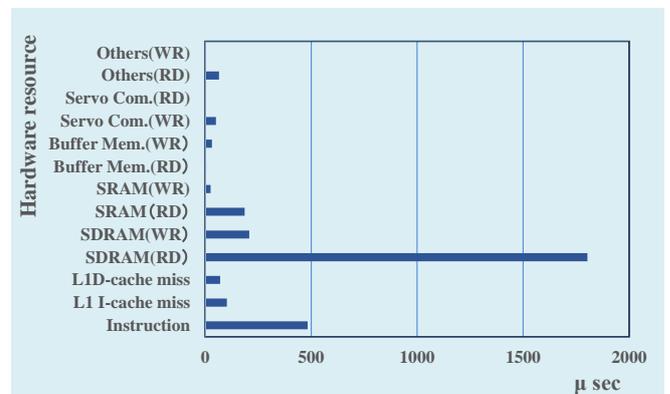


Figure 11: Time Consumed by Hardware Resource.

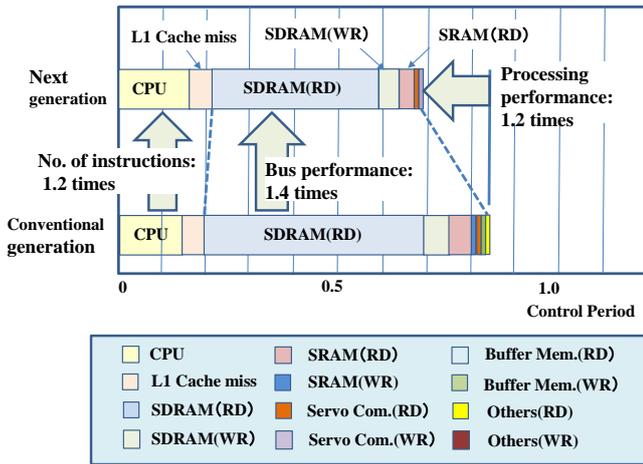


Figure 12: Processing Time Analysis by Hardware Resource

we compared the processing time in the fixed processing cycle time (T) between the conventional system and the system with the proposed hardware architecture. Figure 12 shows the result of analysis of the processing time. The processing time in the fixed control cycle was reduced, and the processing capacity increased by 1.2 times. In order to improve performance of the system, the dominant factor is reduction in the time required for reading from SDRAM. In addition to acceleration of the system bus in line with the increase of the data transfer frequency of the memory, layering of the buses is also effective for communication with the slow devices. By adopting the architecture that combines the buffer RAM and DMA transfer, the system bus traffic also improved.

The bus drive frequency was increased by 1.33 times for the purpose of comparison. By reducing the time required for reading from and writing to SDRAM, layering the memories and buses, and using DMA transfer, the bus processing capacity actually increased by 1.4 times. The number of instructions executed by the CPU increased by 1.2 times, which proves the improvement of the control performance.

5.2.2 System performance verification

To verify the system performance, we created the log information collection program and the aggregation tool, and collected and analyzed logs of the time required to execute each task of the control system. Figure 13 and Figure 14 show the processing time for each of the control tasks shown in Figure 6. Figure 13 shows the comparison between the tasks using the number of execution cycles. In the system to which the proposed approaches were adopted, the processing time was reduced by the improvement of hardware performance for each of the tasks for basic control: synchronization control, pre-processing control, motion control, and machine I/O control. As a result, the processing time distributed to other functions such as operation control or communication control was increased. Therefore, the number of execution cycles increased for "Others".

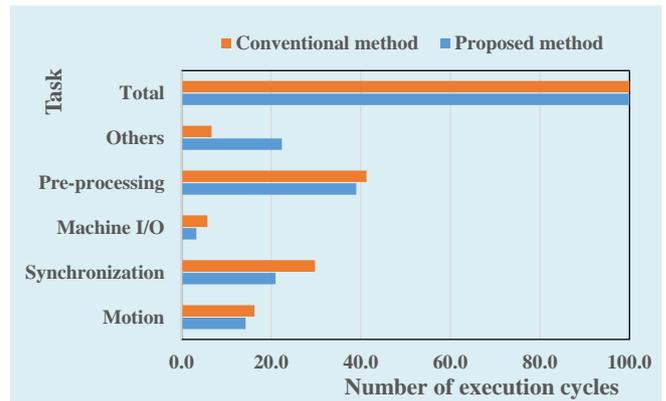


Figure 13: Analysis of the Processing Time by Task Based on the Number of Execution Cycles.

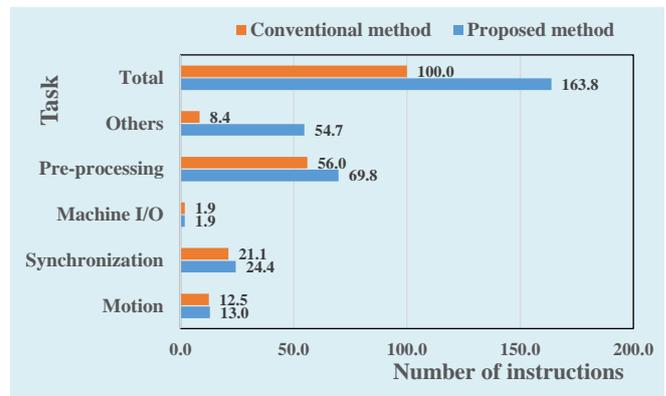


Figure 14: Analysis of the Processing Time by Task Based on the Number of Instructions.

Figure 14 shows the number of instructions for each task in the proposed system, taking the number of instructions executed by the conventional CPU as 100.

For the tasks for synchronization control, pre-processing control, and motion control, the number of instructions executed increased by 20%, which improved the basic performance. The number of instructions executed for other functions also significantly increased. We confirmed that the processing capacity of the system as a whole increased by 1.6 times.

6 CONCLUSION

We reported approaches to enhance the performance of the distributed real-time control system. By the improvement of the hardware architecture by layering of the system bus architecture and the memories, and the use of DMA transfer for the memories and the I/O data, we enhanced performance for basic control and enhanced functions for the system as a whole. Furthermore, we calculated system behavior based on the hardware configuration information and the measured log information, according to which we analyzed and verified the system performance.

It can be said that the approaches to achieve acceleration by improving the architecture by the simulation based on the analysis of system hardware behavior are effective in pursuing higher performance for the distributed real-time

control system. Furthermore, these approaches are essential to achieve acceleration by the system-on-chip (SOC) design. We hope that our report in this paper will contribute to future development of advanced technologies.

REFERENCES

- [1] D. J. Holzmann, U. Mayer, RISC and ASIC-the technologies of the nineties, Proceedings. VLSI and Computer Peripherals. COMPEURO 89, pp.148-150(1989).
- [2] J. Fiddler, E. Stromberg, D. N. Wilner, Software considerations for real-time RISC, Compcon Spring '90. Thirty-Fifth IEEE Computer Society International Conference, pp.274-277(1990)
- [3] John L. Hennessy, and David A. Patterson, Computer Architecture 5th ed., Elsevier, pp. 68-74(2012)
- [4] Jiankang Liu, Yunzhong Fu, Zhenyu Han, Hongya Fu, Design of an industrial Ethernet based embedded open architecture CNC system, ICEDIF, pp.413-417(2015)
- [5] Hongjiang He, Yamin Zhang, Research on Automobile Driving State Real-Time Monitoring System Based on ARM, International Conference on Computer Engineering and Technology, pp.361-364(2009)
- [6] A. Yamashita, H. Mineno, and T. Mizuno, Input/Output Control Method for Serial Communication in the NC Equipment for Machine Tools. Proceedings of International Workshop on Informatics, pp161-167(2015).
- [7] A. Yamashita, H. Mineno, and T. Mizuno, Distributed Remote Input/Output Control Method in Real Time Processing for CNC, International Journal of Informatics Society, Vol.8, No.2, pp67-79(2016).
- [8] Nico Surantha, Astri Maria, Yuhei Nagao, Hiroshi Ochi, Multiprocessor system-on-chip design for industrial wireless application, International Symposium on Electronics and Smart Devices (ISESD) , pp.57-62(2016)
- [9] Fu-Ching Yang, Wen-Kai Huang, Jing-Kun Zhong, Ing-Jer Huang, Automatic Verification of External Interrupt Behaviors for Microprocessor Design, IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, Vol.27, No.9, pp1670-1683(2008)

Session 2:
Intelligent Transportation
Systems
(Chair: Takuya Yoshihiro)

A Traffic Simulation for Improving Traffic Flow in Provincial Cities

Chinatsu Ikeda*, Masashi Saito**, and Ryzo Kiyohara*

*Kanagawa Institute of Technology, Japan

**Kanazawa, Institute of Technology, Japan

{s1421189@cco, kiyohara@ic}.kanagawa-it, msaito@neptune.kanazawa-it).ac.jp

Abstract - Many traffic jams are caused in provincial cities by cars that are used for commuting as well as sightseeing. In many cases, bridges, tunnels, and other such structures become bottlenecks for the traffic. Within provincial cities, there are fewer railways, and therefore public transportation services consist of buses, taxis, and Light Rail Transit (LRT), whose schedule depends on the traffic volume. Therefore, many people cannot avoid the use of passenger cars, which cause traffic jams at intersections, bridges, and tunnels. In order to reduce the traffic volume, many cities are introducing park-and-ride (P&R) systems (an example is shown in Figure 1), road pricing, or toll systems. Road pricing systems, in which drivers have to pay some fee at the entrance gate of a city, have been introduced in some cities such as London, Stockholm, and Milano. However, these are large cities that already have many public transportation services, such as subways or monorails. However, even after introducing P&R systems, heavy traffic jams still occur. In this paper, we focus on the P&R system and discuss issues that cause heavy traffic jams. Then, we propose and discuss a new incentive-based P&R system.

Keywords: ITS, Traffic jam, Simulation, Park & Ride, Open data.

1 INTRODUCTION

Traffic congestion in urban areas is getting worse. The traffic congestion in many large cities can be avoided by increasing the use of public transportation services, for example, when commuting to school. In future, autonomous vehicles can be used by people who cannot drive.

The speed of autonomous vehicles would be low during the initial stage for safety reasons, and it is expected that the volume of traffic that can be passed in a certain time will be reduced. On the other hand, there are opinions that the traffic congestion would be reduced by efficient driving.

In urban areas such as Tokyo, there is sufficient financial prowess to construct new roads or means of underground transport. Moreover, in urban areas, the public transportation systems are highly developed; therefore, there are few problems in such cities. However, in many smaller cities (not just in Japan), the public transportation systems are not well developed, and the means of commuting are limited. Therefore, many people use vehicles, which cause traffic jams. Therefore, the reduction of traffic congestion in smaller cities is an important research theme.

In provincial cities, there are cases in which even when new roads are constructed, few vehicles use the new roads. In many cases, this is due to incorrect simulation. Therefore,

effective and effective measures are required to solve such issues. Thus, accurately performing advance simulations is an important issue. In order to implement policies and reduce traffic congestion, it is necessary to clarify the purpose and other such details of a passing vehicle. For example, in the following cases, it should be considered that the number of vehicles would not decrease no matter how many measures are taken.

1. In the case of lack of public transportation or any other means of transportation other than the vehicle.
2. In the case of absence of any option to carry objects such as large baggage, which cannot be held by hand.
3. In the case of truck or taxi.

On the other hand, it is considered possible to use public transportation for commuting to and from work, instead of driving.

In order to make these possible, various methods have been introduced in many places such as the park & ride shown in Fig. 1 and road pricing using the model shown in Table 1. In Japan, traffic demand management [1] has been promoted since the 1990s, and P & R systems etc. are widely practiced in many places in Japan such as Sapporo City in Hokkaido and Kanazawa City in Ishikawa Prefecture [2].

Therefore, new measures are necessary; however, they are

Table 1

Type	Target	Implementation
Fixed billing	Fixed Area	Tollgate for fixed area
Dynamic Billing	Fixed Road	Tollgate for fixed road

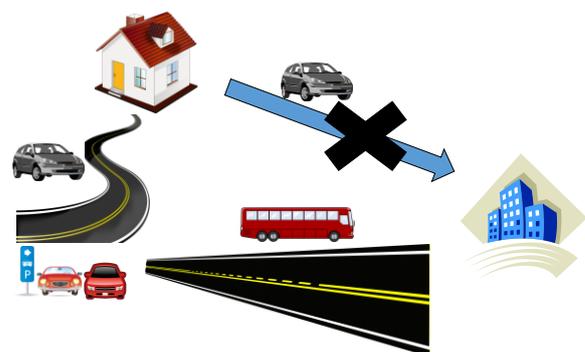


Figure 1. Park & Ride

Table 2 Items in Mobile Spatial Statistics

area	Date	Day of week	Time	Area code	Residence code	age	gender	population
------	------	-------------	------	-----------	----------------	-----	--------	------------

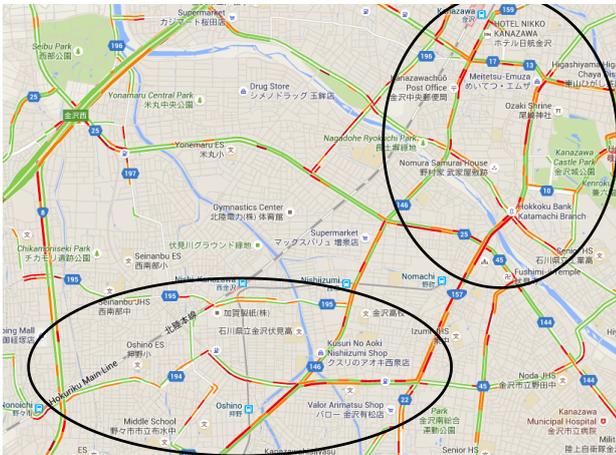


Figure 2 Kanazawa City & Nonoichi City

necessary to perform the simulation accurately beforehand. For this purpose, it is necessary to know the drivers' OD (Origin Destination) traffic volume [3]. In the conventional method, personal trip surveys are used [4]. However, it is difficult to use personal trip surveys to obtain the movements of people, which change each moment depending on the time of day. Moreover, this depends on people's memory, which makes it difficult to obtain information such as the correct time, etc.

On the other hand, details of the investigation of traffic volume "traffic census", which considers only the vehicles on the road, is made public once every five years. This method involves manually counting the number of vehicles at intersections and similar locations. In addition, at present, "Mobile Spatial Statistics" are tabulated as shown in Table 2. This discloses information on the terminals in a cellular base station as open data.

We propose a traffic volume forecasting method using traffic data from the "traffic census" survey on traffic volume, which is available once every five years. We have already studied the method [5, 6]. Based on the results of these studies, we apply the prediction method based on the traffic survey data and the number of times a mobile phone user stays in the coverage area of a base station by using the example of Kanazawa city in Ishikawa Prefecture and its surrounding areas shown in Fig. 2.

2 RELATED WORKS

Several methods of forecasting traffic have been proposed. In particular, there is a method that utilizes personal trip information. However, this method has problems related to the cost and accuracy of obtaining personal trip information.

Meanwhile, a method of extracting the trip information of a person through a smartphone application has been proposed [7]. However, this is related to the method of

acquiring personal trip data, and it cannot be used immediately because of the necessity of installing an application.

The method proposed in [8] is too limited and cannot be extended to the behavior of ordinary people. Several methods that utilize the base station information of mobile phones have been proposed. These methods are based on the premise of using open data published as mobile space statistics of Kashiwa City and Yokohama City. [9] In this study, a regional evaluation model was studied to analyze the feasibility of using mobile space statistics. The regional characteristics show that it is possible to evaluate areas that were not covered in previous statistics by using the population attribute during daytime. In comparison with geographical characteristics, mobile space statistics have been found to possess more relevance to regional evaluation. Thus, Mobile Spatial Statistics is considered as an effective tool for regional assessment.

However, from these characteristics, it is impossible to estimate the actual moving population without knowing the current population of the area.

There has been a research on obtaining personal trip data from base station information instead of mobile space statistics [10]. In this method, by using the IDs of cellular phones in the coverage area of a base station, various details are obtained from the relationship between the distance and time, based on where the ID would be at the next instant. However, in general it is difficult to link IDs that become personal information, and it is not easy to use this method.

A method of estimating the OD traffic volume based on the transportation method using the staying population data has been proposed [11]. In this research, the authors have proposed and implemented a method for the introduction and evaluation of a new public transportation system "Smart Access Vehicle (SAV)," which makes no distinction between a bus and a taxi, owing to the ease of obtaining the observation data, the number of people staying in a zone (from mobile zone statistics), and the number of passengers on specific bus routes. We propose an estimation method for OD traffic volume based on the transportation method, using the number and small traveling locus data of a small sample. To verify the accuracy of this method, we calculated the traffic volume of 200 zones of 500 m mesh in Hakodate city. The model for selecting the transportation mode was created, the model parameters were estimated, the service level was set, and the simulation was carried out.

We estimate the number of users by the mode of transportation; however, since it is not based on precision problems or automobile traffic volume survey data, the actual traffic volume cannot be predicted, and it is cited as a future issue.

In this study, we propose and evaluate a method to predict the personal trip information and the number of travelling vehicles (without conducting a personal trip survey),

confirm the current state via simulation, and predict the impact of P & R system.

3 PROPOSED METHOD

It is necessary to predict basic personal trip information and the means of transportation from the transition of the number of terminals existing in the base station of the mobile phone using the mobile space statistics, which is shown in Fig. 3, and the measurement result of the traffic volume.

3.1 Mobile Spatial Statistics

The "Mobile space statistics" were provided by NTT DoCoMo, Inc. [9], and it includes periodical data on the mobile phones that exist in the area of each base station. It is assumed that the population of daytime and nighttime can be predicted and that the number of people that have moved to the city center at a particular time can be obtained.

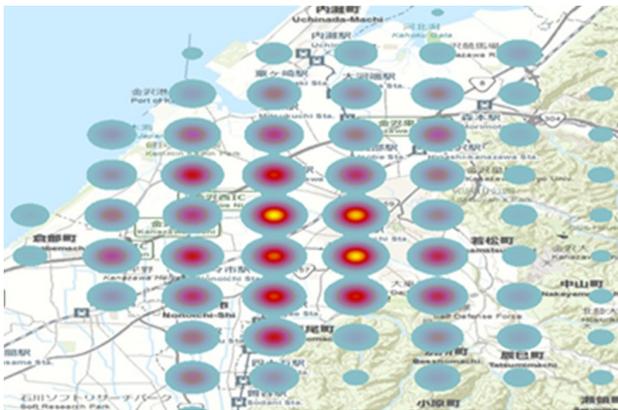


Figure 3 Example of distribution of terminals (Google Map)



Figure 4 Exapmle of measured point of traffic (Google Map)

The data for each place of residence of Kanazawa city, Nonoichi city, Hakusan city, and Ishikawa prefecture were used for the study. The information contained in the "Mobile Spatial Statistics" includes date, time, area, place of residence, age, gender, population, and this information is based on cell phone contracts. If the phones have been bought by parents for their children, the parents' names would be registered at the mobile company, and the details such as age would be that of the parent. However, this information is sufficient to know the overall trend.

The "population" field indicates the number of terminals. The actual area is divided into a mesh, and the area is confirmed. The age is based on information obtained from the contracts. It is categorized into the following groups: 15 to 19 years, 20 to 29 years, 30 to 39 years, 40 to 49 years, 50 to 59 years, 60 to 69 years, and 70 to 79 years. However, in the lower and older age groups the contracts may not necessarily be in the person's name. It is also assumed that the registration of family members is not limited. With respect to the problem of personal information, processing is done so that the details of an individual are not specified, and since the names on the contract data is removed, it is safe. Table 2 shows an example of this data.

3.2 Road Traffic Census

The Road Traffic Census [12] is used by the Ministry of Land, Infrastructure, Transport and Tourism to investigate the nationwide road situation, traffic volume, travel speed, departure place/destination of automobiles, purpose of operation, etc. once every five years. The contents include the traffic volume between roads for 12 hours, the traffic volume for 24 hours, the average traveling speed at the time of congestion (during peak times in the morning or evening), and the average traveling speed at the time of non-congestion. In this study, the latest data from the Road Traffic Census is used. Although this data is 7 years old (from 2010) it seems that there will be no problem in using this data. The latest version is scheduled to be released in 2017, and once it is released, we will update this with the new data as soon as possible. An example of the survey points in Heisei 22nd edition is shown in Fig. 4, and an example of the survey data is shown in Table 3.

3.3 Visualization of Mobile Spatial Statistics

By visualizing the mobile spatial statistics data, we can see the movement of people, including the area from which people are moving the area to which they are moving. From this difference in movement, it can be estimated that this difference is likely due to commute. In addition, it is assumed that the actual number of people can be obtained by estimating the average number of families that are away from the city center and the actual number of workers

Table 3 items of measured

Road	Traffic volume of date time	Traffic volume of a day	Average velocity of crowded situation	Average velocity of non-crowded situation
------	-----------------------------	-------------------------	---------------------------------------	---

Table 4 assumed parameter

Traffic volume of each area	a
Traffic volume of commuter	s
People who have to use vehicles	D

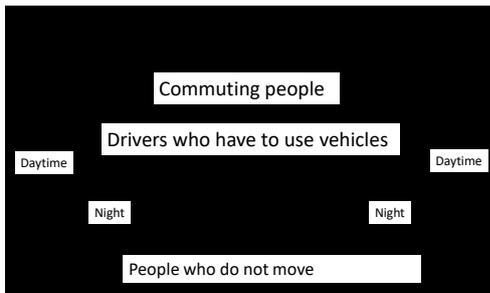


Figure 5 Images of estimation



Figure 6 mobility of terminals

among them. Assuming that most of them commute by car, we obtain the number of vehicles that are not moving (unless it is a car) by subtracting the traffic volume of the corresponding commuter from the traffic volume in each area. The number of vehicles thus obtained becomes the traffic volume, which cannot be reduced by P & R and road pricing.

These assumptions are shown in Table 4. D is obtained from the following equation.

$$D = a - s \quad (1)$$

Conversely, with regard to the traffic volume that can be reduced, it is estimated that the number of major parking lots in the area and the number of vehicles are reduced, when the maximum volume is obtained from the information of public transportation facilities. The estimated image is shown in Fig. 5. By visualizing the mobile space statistical data (Fig. 6), it is possible to obtain information on the movement of a person, and from the increase/decrease in difference of each area, it can be seen

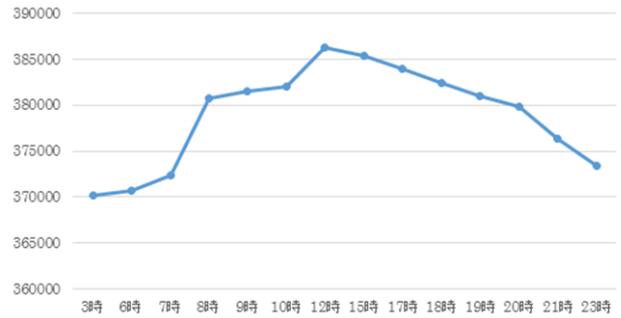


Figure 7 population of each time



Figure 8 Average moving of each time

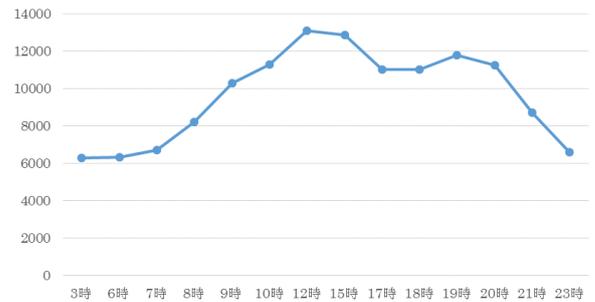


Figure 9 Vehicle volume of each time

from which area and to which area the person moves. Moreover, from the difference in day and night population, it is seen that the possibility of commuting is large. Assuming that many of them are commuting by car, we estimate the number of vehicles by subtracting the traffic volume of the corresponding commuters from the traffic volume in each area.

3.4 Estimation of peak hour and number of vehicles

we estimate the number of moving vehicles from the mobile space statistical data of the peak time and the population difference. We estimate the peak time and use the population difference between day and night to predict the number of people who are supposed to be traveling by car etc. Fig. 7 shows the population by time, and Fig. 8 shows the number of moving vehicles. To estimate the number of moving vehicles, the population is estimated by the incremental/decremental difference in the number of terminals of time events in each area. From the two figures, it can be seen that the largest number of moving vehicles

Table 5 Kpark and capacity

Park & Ride	volume	Bus stop
Osaka-ya	5	Chikaoka
Genki	38	Kan-non-do
Maxvalue	5	Kitamachi
Maruei	5	Magira-2
eaon	20	Wakamatsu

was during the peak time around 7:00 and that about 8298 people were moving. The population was the largest at around 12:00. In the morning, the number of movements is large because there are many people who commute or go to school, and the increase in the number of movements again around 10:00 is mainly due to people who ride a car to work.

Moreover, from the difference in population between day and night, it can be seen that the number of commuting schoolchildren was about 12,851. By using equation (1), it can be estimated that 6,729 people out of about 19571 people are traveling by car etc. in the area around Kanazawa station at 7:00. Since these 670 people do not have to be in a car, depending on the measure, it is a layer in which public transport can be used.

According to the Road Traffic Census, the place with the most traffic volume in Kanazawa City was Xi'an, where the traffic volume (for cars) during daytime was about 61,000 units,. However, from the statistical data, the traffic volume for Xi'ni was calculated to be 62,341 units. This error is considered to be caused because there are several people in one car and there were several units outside the observation point of the road traffic census. Fig. 9 shows the variation in the number of units in all areas. In Figs. 7 and 8, a variation can be seen in the morning ; however a small difference was seen from the afternoon. It seems that this is influenced by differences in the date of the survey as well.

4 SIMULATION

4.1 Current Status of P & R and an Example of Improved Method

Since 1996, a system called "K park" has been implemented as a countermeasure against traffic congestion in Kanazawa city of Ishikawa Prefecture. This is a P & R system targeting commuters and schoolchildren in cars. In addition, in order to make it easy to use, we have introduced a parking guidance system since 2001, so that the state of the parking lot in each district can be seen from the homepage of the parking guidance application. Table 5 lists the parking lots of K park and the number of parked cars.

The peak traffic situation is shown in Fig. 9, and the traffic volume from the road traffic census is shown in Table 3. It can be seen that the number of vehicles that can be parked is obviously small for the traffic volume. When the number of parking lots is small, the driver needs to consider measures to be taken when the parking lot is full, and the

Table 6 Road and Capacity of parking space

Road	Number of vehicles
1	1500
2	1200
3	1000
4	1300

Table 7 Setting of simulation

Human agents	61000 people
Vehicle agents	61000 vehicles
Number of passengers	1 person
Velocity of vehicles	42km/h
Seating capacity of a bus	30 seats
Service frequency	30 minutes
Velocity of people	4km/h
Number of lanes	2 lanes
signals	12 signals
intersections	8 intersections

trouble of searching increases even if the total capacity is sufficient. This may be preventing the spread of P & R system.

Therefore, by assuming to use a parking lot of a large shopping center, we will evaluate via simulation what will happen if the number of P & R users increase.

4.2 Experiment

In this experiment, we used Scenargie [13], a multi agent traffic simulator developed by Space-Time Engineering Company. The scenario considers the time from 6:00 to noon. All the human agents move to the designated parking lot, walk from there to the bus stop, take a bus, and move.

Experiments were conducted on two patterns of increased parking capacity: the case when the capacity of the current parking lot was increased, and the case when a large parking lot was used (for this, a large parking lot of a large shopping center located in the suburbs of Kanazawa city was used).

The parking lot extends to the front of the station, and the large parking lot of the store is located near the road. Fig. 10 shows the simulation points, and Fig. 11 shows the map used for simulation (map published on OpenStreetMap [14]). The departure point is from the observation point to the front of the station. The vehicle agent moves to the four places around the station on the same route, and all the roads have two lanes on each side. Overtaking is over, and there is a lane change accompanying the right/left turn, heading off. Table 5 shows the number of cars that can be parked at the large parking lot of the store.

The number of parking lots is assumed by considering the parking lot of a shopping mall with a large number of cars parked near the road. Table 6 and Table 7 list the simulation settings. The number of human agents and vehicle agents, and the speed of the car are determined using the 12-hour car traffic volume from the road traffic census and the



Figure 10 Simulation points

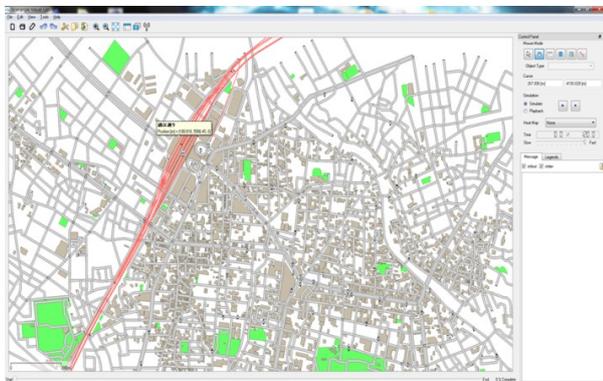


Figure 11 maps of simulation (Open street map)

average travel speed. We did not set any fee for bus etc.; we made a default except for the car and person's moving speed.

4.3 Result

Fig. 12 shows the result of P&R using the large parking lot of store. Fig. 13 shows the result of increasing the number of P&R available for cars in the current situation. Table 8 and Table 9 show the reducing ratio of using large parking spaces and additional parking spaces, respectively.

Compared with the case of increasing the large number, the effect was obtained throughout the road. There is a parking lot and a bus stop near the store with the large parking lot, and the increase in the number of vehicles that can be parked are considered to be caused by the difference in the distance between the parking lot and the bus stop, which depends on the road. Regarding the difference in effect depending on the roads, on road 1, the parking capacity is 1,500 and on road 2, it is 1,300. The traffic congestion and the number of vehicles that can be parked may have occurred due to the parking capacity and the number of right or left turns or signals in the road. As for the increase, it is considered that the difference is due to the

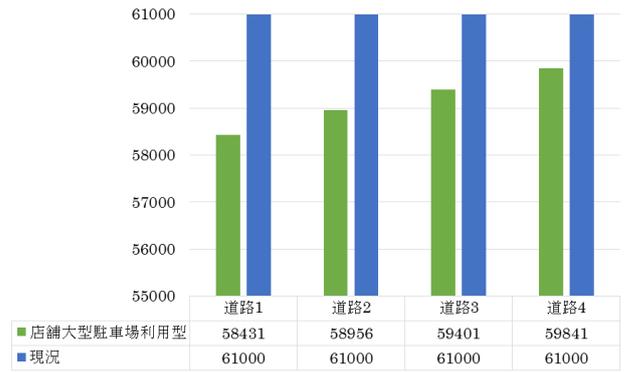


Figure 12 Result of using large parking spaces

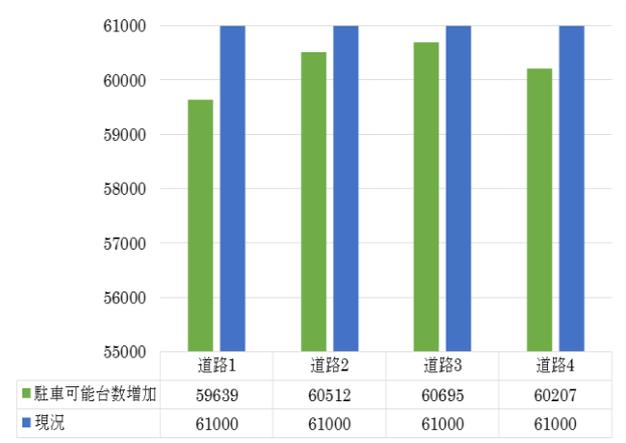


Figure 13 Result of additional parking spaces

Table 8 Reducing ratio of using large parking spaces

Traffic Volume	Current Status	Using shopping center	Reducing ratio
Road1	61000	58431	4.21%
Road2	61000	58956	3.35%
Road3	61000	59401	2.62%
Road4	61000	59814	1.90%

Table 9 Reducing ratio of additional parking spaces

Traffic Volume	Current Status	Extension of K park	Reducing ratio
Road1	61000	59639	2.23%
Road2	61000	60512	0.80%
Road3	61000	60695	0.50%
Road4	61000	60207	1.30%

traffic congestion caused by the concentration of parking lots in front of the station.

In section 3.4, it was shown that the current population is headed towards the station and is concentrated. As a result of simulating the kinds of measures that can be taken to reduce traffic volume, there are many shopping malls in Kanazawa city and its surroundings. In many cases, they possess large parking lots. In the current situation, there are bus stops near some large stores, and there shuttle buses every hour. In addition, if we increase the number of P & R

available for cars, there will be a change in the traffic volume. If the number of parking lots is small, we have to consider measures to be taken when the parking lot is full. Even if it is enough, it will increase the trouble of searching. Increasing the number of parking lots increases the number of parking lots in the parking lot so the distance to the bus stop has increased. Furthermore, in this simulation, it can be considered that the waiting time increased because there was only one bus stop near the parking lot. In order to reduce the traffic volume by P & R, we found that it is necessary to review the number of available parking lots and the location of the bus stops, and to increase the number of buses being operated.

5 CONCLUSION

In this paper, we proposed a traffic forecasting method based on mobile space statistical data and road traffic census for Kanazawa city, Ishikawa prefecture. In the proposed method, it is possible to obtain the number of moving population from mobile space statistics data. Further, the number of commuting school students is estimated from the moving population, and it is presumed that they are traveling by car, etc. In future, it is necessary to develop traffic volume forecasting models that incorporate not only road traffic census but also other road information. Simulation was conducted on the effect of increasing the number P&R using the parking lot of a large-scale store, taking advantage of the characteristics of the target city and by increasing the available parking capacity of P&Rs. Since the road traffic census data used in this study was for FY 2010, the latest general traffic-volume survey data is needed.

In addition, although mobile spatial statistical data was used for each municipality, it is thought that finer data on the movement of people can be obtained by combining data related to gender and age as well. Regarding the simulation, it is necessary to set up a model for public transportation and a behavior model with detailed parameters. In addition, simulations that consider transportation systems such as taxis, railway, and bus are necessary. From the statistical data, it is possible to estimate the moving population and calculate the number of commuting commuters from among them, and show that the population considered to be traveling by cars etc. can be estimated. Moreover, through the simulation, it is shown that it is necessary to review the number of vehicles that can be parked and the position of bus stops, and to increase the number of buses being operated.

ACKNOWLEDGEMENT

This work was supported by JSPS KAKENHI Grant number 16K00433

REFERENCES

- [1] K.Asai, "Dictionary for road and way," NIPPON JITSUGYO PUBLISHING, 2001 (Japanese)
- [2] S. Aoshima, S Suda, et al. "Analysis on Cost-Time Characteristics of Park & Ride Trip and Necessary Conditions of the Parking in Provincial Regions," INFRASTRUCTURE PLANNING REVIEW, Vol. 16 (1999) P 863-868 (Japanese)
- [3] R.Kanamori, K.Mizuno, I. Noda, H. Nakashima," Estimation of Multi-Mode Origin-Destination Flows by using Staying Population Data at Base-Stations," IPSJ 2015-ICS-178 (Japanese)
- [4] The Ministry of Land, Infrastructure, Transport and Tourism, "Report of the Fifth Person Trip Survey ," 2009 (Japanese)
- [5] K. Enomoto, M. Saito, R. Kiyohara," Prediction of Traffic Flow in Provincial Cities by Open Data," IPSJ 2016-MBL-80, 2016 (Japanese)
- [6] K. Enomoto ,M. Saito ,R. Kiyohara, "A Study of Park-and-Ride Systems in Provincial Cities," The Ninth International Conference on Mobile Computing and Ubiquitous Networking(ICMU2016)
- [7] H.Nakamura, K. Miyashita, E. Hato, T. Kishii," The Estimation Method of Transportation mode for Person Trip Survey using Acceleration Sensor with Random Forest," JSTE Journal of Traffic Engineering Vol. 1 (2015) No. 5 p. 10-18 (Japanese)
- [8] H. Kobayashi, S. Hirata," A Proposal of Evaluation Method for People with Mobility Difficulties using the Person Trip Survey Results," Journal of the City Planning Institute of Japan, Vol. 48 (2013) No. 3 p. 159-164 (Japanese)
- [9] T. Seike, H. Mimaki, S. Morita, " RESEARCH ON THE EVALUATION OF REGIONAL PECULIARITIES IN KASHIWA AND YOKOHAMA BY "MOBILE SPATIAL STATISTICS", " AIJ Journal of Technology and Design,Vol. 21 (2015) No. 48 p. 821-826(Japanese)
- [10] Y. Yamada, A. Uchiyama, A. Hiromori, H. Yamaguchi, T. Higashino, "Proposal of a Travel Estimation Method Using Control Signal Records in Cellular Networks and Geographical Information, IPSJ Journal, 2016-08-15, pp. 1826 – 1834, (Japanese)
- [11] H. Nakashima, S. Sano, K. Hirata, Y. Shiraishi, H. Matsubara, R. Kanamari, H. Koshiha, I. Noda, "One Cycle of Smart Access Vehicle Service," ICServ2014
- [12] The Ministry of Land, Infrastructure, Transport and Tourism, "Census of Traffic Volume" <http://www.mlit.go.jp/road/census/h22-1/>
- [13] Scenargie, <https://www.spacetime-eng.com/jp>
- [14] Open Street Map Japan, <https://openstreetmap.jp/>.

A Study of Driver's Contexts for visibility of equipment on Vehicle

Yuki Chigira*, Yuichi Tokunaga** and Ryozo Kiyohara*†

*Kanagawa Institute of Technology, Japan

**Mitsubishi Electric Corp., Japan

{kiyohara@ic, s1421099@cco}.kanagawa-it.ac.jp

Abstract - Location-based services (LBSs) have emerged as a popular means of providing services using location data. For example, providing services based on location data that can be easily accessed through smartphone applications can ensure more convenient and stress-free driving experience. To provide satisfying services, LBSs should refer to the users' contexts as well as the location data, particularly when the context represents the relation between the user and a specific location. This study proposed a method to estimate the location recognition level at a specific location.

The proposed method is based on the processes associated with the construction of cognitive maps. It uses location and landmark data, which significantly impact cognitive mapping building processes. Here, it is deemed desirable to estimate the location recognition level using smartphones, which are always carried by users, together with the LBS platform. Therefore, the proposed method relies only on location and landmark data, which can be gathered by smartphones. In addition, experiments were conducted to verify and evaluate the proposed method. The experiment uses driving videos to measure the location recognition level of the 12 participants. The result shows that an accuracy of approximately 50% can be achieved when the proposed method is used to estimate the recognition level of a location.

Keywords: ITS, Context Aware, LBS, Cognitive MAP, Vehicle Information Devices.

1 INTRODUCTION

Attention has been paid to the development of autonomous vehicles that could improve road safety by reducing accidents caused by the increasing number of senior drivers, reduce traffic congestions, and CO2 emissions. It is anticipated that the autonomous vehicle technology will gradually be enhanced and put into the market. The National Highway Traffic Safety Administration (NHTSA) identifies five levels of automated-driving in which Level 5 refers to a fully autonomous system with no driver needed to operate the vehicle.

In order to enhance the driver's safety, even level 5 automated systems should offer the driver the necessary information about the vehicle situation and performance. Thus, it is important to provide the driver with the necessary information in an easy-to-understand manner based on the contextual driving conditions including the driver situation, vehicle situation, and surrounding conditions.

Drivers are increasingly using smartphones as navigation tools and audio devices during driving [2]. The various

contextual driving conditions, which include driving contexts (running, stopping, fast moving etc.), surrounding context (weather, traffic congestion, etc.), and driver's context (drowsiness, fatigue, disorientation or lostness, etc.) are constantly changing, and the service required by the user in operation continues to change accordingly. Currently, there are few location-based services (LBSs) other than maps and navigation systems that can be used by drivers. However, as the number of drivers who use smartphones as in-vehicle information terminals is increasing, the position context of the driver becomes important because LBSs can be used to deliver contextually aware services.

The previous experimental studies have confirmed the importance of driving context recognition by examining the composition process of the driver recognition map in terms of recognizing the number of roads and landmarks passed by the driver [3-5]. This study further investigates this issue by evaluating the drivers' ability to recognize the differences between day and night and the type of landmark.

2 RELATED WORKS

Many studies were conducted on context recognition during driving. For example, Han et al. focused on dozing detection [6], Yang et al. estimated the drivers' degree of fatigue [7], Ji and Yang assessed the drivers' alertness and awareness of their surroundings and determined whether safe driving was achieved [8]. Many studies focused on recognizing the drivers' reactions to maintain safe conditions [9]. These contexts are very critical to address safety and usability issues.

The location recognition level addressed in this study is not directly related to safety improvement such as detection of dozing or estimation of fatigue degree; however, providing services based on the drivers degree of location recognition can improve safety and contribute to improving usability.

3 OUTLINE OF COGNITIVE MAP AND DEFINITION OF LOCATION RECOGNITION LEVEL

3.1 OUTLINE OF COGNITIVE MAPS

A cognitive map is a person's mental image that represents the spatial knowledge of physical locations and

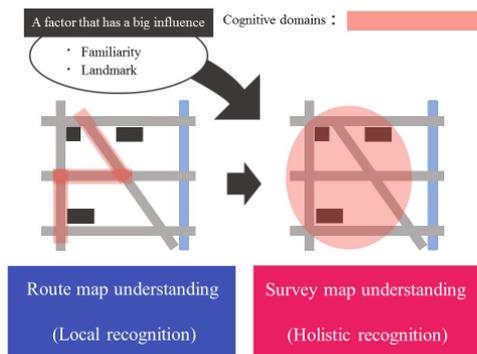


Figure 1 Construction process of cognitive map

allows visualizing images of spaces. The state of how much the user recognizes a specific place can be replaced with the state of how clearly a specific space is constructed in the cognitive map held by the user. In other words, the clearer the image of this space, the higher the location recognition level. Therefore, understanding the construction of peoples' cognitive maps can allow estimating their level of location recognition.

The research on cognitive maps was initially developed by Tolman when he first conducted experiments using rats [10]. The experiments involved releasing rats in a maze where food was placed at the exit. It was found that as the experiment was repeated, the number of times the mice entered the blind alley was reduced. Tolman analyzed that the rats learned space by exploring the maze and built mental images of the space. Such images were defined as cognitive maps. Images of space are thought to be similarly held by humans. Research studies were conducted from various viewpoints in the field of cognitive psychology and applied psychology to clarify the relationship between humans and space. Cognitive maps can be useful for elucidating and analyzing human spatial behavior and design of cities [11, 12].

It has been reported that cognitive mapping involves two stages, namely "cognitive map understanding" and "survey map-like understanding" [13]. Fig. 1 shows the outline of the process of building a cognitive map. At the route map understanding stage, the people spatial cognition of the route viewed from their own point of view and their surroundings is performed within the local region. At the stage of survey map understanding, a bird's-eye view of spatial perception is performed covering a wide area. In other words, the route map understanding stage only involves developing a procedural image of a moving experience, such as "go right at the intersection and then straight ahead to the next signal" without developing positional relationship with the surrounding space. However, as the construction of cognitive maps progresses and reaches the stage of survey map understanding, the positional relationship between the spaces can be recognized; thus, they become capable of directing themselves or to select a suitable route for their partner.

Two factors have been reported to influence the process of cognitive mapping moving from route map understanding to survey map-like understanding, namely "familiarity" and

"landmark." Based on the result of the study on verifying the cognitive map of urban residents conducted in Venezuela considering the number of years of residence, Appleyard focused on the elements captured when learning a space that was not familiar, and analyzed how the perception of the space changed as people became familiar with the space [14]. Therefore, the number of running times can be used as a factor representing familiarity with the space to estimate the level of the driver location recognition.

The research related to the influence of landmarks revealed that landmarks are typically stored since the initial stage of developing the cognitive map [15]. Additionally, the spatial information (direction and distance) and the perceptual information of such landmarks are often used [16]. The number of landmarks captured increases in proportion to the detailed cognitive map developed by the person [17]. It is reported that the number of landmarks that can be stored is limited to the short-term storage process of working memory that can hold only the magic number of 7 ± 2 [18, 19]. In other words, landmarks have very important role in the process of increasing the human level of location recognition. Hence, landmarks are considered as important reference information when estimating the location recognition level.

Based on the above discussion, two factors are most likely to influence measuring the drivers location recognition level, namely 1) familiarity and 2) landmark.

3.2 CONFIRMATION EXPERIMENT

Many of the studies related to cognitive mapping mentioned in the previous section were dedicated to investigating pedestrians' navigation behavior, and there is limited research on driver's cognitive mapping. Considering that pedestrians and drivers have different moving speeds and subjects to grab their attention during traveling, the factors affecting the process of building a cognitive map may also differ. There is some research in the field such as the study by Kubota et al. [20] that examined the cognitive mapping of drivers. The study findings revealed that the density of information (the number of buildings and other features recalled after passing the same route) is similar between pedestrians and drivers, despite the difference of the cognition content.

Therefore, before developing a method to estimate the place recognition level using information such as familiarity and landmark, it is necessary to confirm that such information actually influences the process of building the driver's cognitive map. The previously conducted experimental studies [3-5] confirmed that landmarks have an influence on the construction of the driver's cognitive map.

3.3 DEFINITION OF LOCATION RECOGNITION LEVEL

Location recognition level can be defined based on the above-mentioned cognitive map construction process. Three stages of location recognition are identified in this study, namely "understanding well," "recognizing to a certain

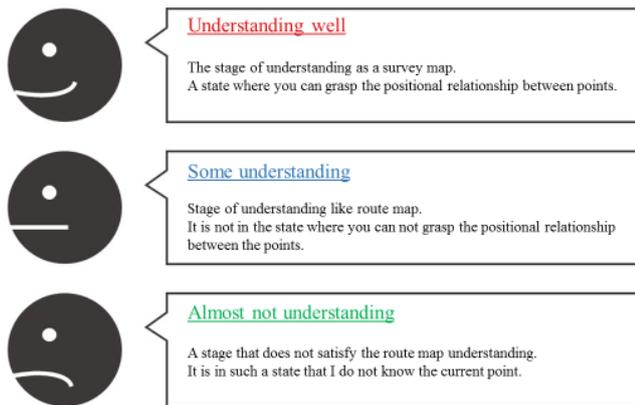


Figure. 2 Overview of definition of point grasping degree

extent,” and “hardly recognizing” the place. Fig. 2 shows an illustration of these stages, and each stage is defined below.

■ Understanding well

It is defined as a stage of a survey map-like understanding in the process of building a cognitive map. This stage does not require external resources such as maps for movement, and can also select and guide the appropriate route for the partner (cultivate new route).

■ Have some recognition

It is defined as a stage of route map understanding in the process of building a cognitive map. This stage does not require external resources such as maps for movement, as it does not involve developing a bird's-eye view image to recognize the surrounding locations. However, it is necessary to be able to perform the routing and select an appropriate route for the partner.

■ Do not have much recognition

It is defined as a stage that does not satisfy the route map understanding in the process of building a cognitive map. This stage requires external resources such as maps for movement because the person is almost in the state of first looking at the place.

4 PROPOSED METHOD

4.1 OUTLINE OF PROPOSED METHOD

As mentioned in Section 3, the cognitive map (mental image of the space) held by a person is deepened as it is influenced by familiarity and landmarks. As mentioned above, not only pedestrians but also drivers are similarly affected by these factors. By repeatedly visiting a certain place, the recognition level of the place becomes high, and the recognition degree of the places that have never been visited is low. This demonstrates the influence of familiarity. Even when occasionally visiting a place with buildings that leave impression, the recognition of the place will be high. This shows the influence of landmarks. These effects will be quantitatively measured.

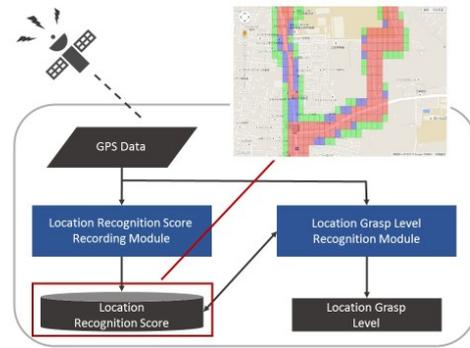


Figure 3 . Proposed method

There are further factors to be considered including both the temporal and spatial factors that can affect the degree of detail of the developed cognitive maps. Such factors include the changes in the surrounding environment such as visiting during the day or at night and the time elapsed since visiting the site last time (familiarity interval).

Therefore, an approach is adopted to examine the cognitive map development processes by quantitatively measuring these factors. In this approach, the travel frequency is considered as familiarity. The number of times a person passes by a certain place in the past can be used to quantitatively measure familiarity. The travel frequency information can be obtained by recording the position information using the smartphone. Meanwhile, the landmark location information can be obtained via the Internet. The influence of landmarks on the user can be quantitatively measured based on whether the landmark exists and the landmark impression is captured by the driver.

Cognitive maps that are quantitatively processed are also preserved in the past state. They are updated every time the person experiences space. In this manner, the recognition level of each place is recorded as a score by using the position information and the landmark information. The updating process every time the space is experienced is performed by the “location recognition score recording module.”

The proposed method consists of “location recognition score recording module” and “location recognition level module.” The overall structure of the proposed method is shown in Fig. 3.

4.2 LOCATION RECOGNITION SCORE RECORDING MODULE

In order to quantitatively examine the cognitive map held by the user, the location recognition score recording module records the recognition degree at each location as a score by using the position information and the landmark information. The module updates the recognition degree each time space is experienced. The outline of this module is shown in Fig. 3.

The location recognition score is recorded in lattice-like latitude/longitude coordinates (mass) divided at a fixed distance. As the user cognitive map is updated every time he/she experiences a space, the past recognition score is also

recorded in past values, and updates are made each time the space is experienced. Locations that were not recorded in the past are given 0.0 value.

In this study, each square lattice was 25 m square. More details are described in Section 4.3. To calculate the location recognition score, a filter of 33 square centered on the position of the user is used. The value of the filter will be described later. The value of each filter square is calculated in consideration of the surrounding environment (daytime or night) and the influence of landmark. The sum of the values of 9 squares is added to the score recorded at the point where the user passed.

The processing procedures of this module are summarized as follows:

1. Acquire the location information of the point the user passed.
2. Prepare a filter of 33 square centered on the point through which the user passed.
3. Calculate the value of each square in the filter considering the influence of landmarks and surrounding environment (daytime or night).
4. Add up the values of all filters in the filter.
5. Acquire the location recognition score already recorded at the point where the user passed.
6. Add the total value to the point recognition score already recorded.
7. Record the added value at the point where the user passed.

4.2.1 FILTER VALUE

The value of the filter used for recording the location recognition score is 1) the front of the user, 2) the front left and right of the user, the position of the user (mass of the center), 3) the value of the filter used for recording the backward location recognition score, and 4) weighted in order of the backward left and right of the user, so that the sum is 1.0. The specific value is the "initial value of filter" shown in Fig. 4.

Different filter values must be used for daytime and nighttime because the visibility decreases markedly at night, as shown in Fig. 5. A study has been conducted to verify the difference in the visibility level when a car traveling at night with a high beam and when it is set to a low beam. According to the study, the high beam has a safety visible distance of 2 to 3 times that of the low beam. This study compared the safety visibility distance between the high beam and the low beam; however, it can serve as a reference to quantitatively measure the visibility difference between daytime and nighttime. Even with high beam, the visibility is inferior to that in the daytime. Therefore, based on the values obtained in this report, the value of all the cells of the filter divided by 4 is considered as the nighttime filter. Although the visibility at night varies between the locations with streetlights and those where there are no streetlights, for simplicity, the value obtained by dividing the filter value by 4 at all points is used in this study.

In addition, in order to quantitatively measure the influence of landmarks on the process of constructing cognitive maps, the information on whether landmarks exist

and the landmark impression that was captured by users is evaluated, and the obtained filter values are changed accordingly. When there is a landmark in the applicable range of the filter, the value of the filter is changed according to the following procedure:

1. Check whether a landmark exists in the area corresponding to each square of the filter.
2. If there is a landmark, multiply the value of the corresponding square by the weight of the landmark, as will be described later.

When there are multiple landmarks in a square, the largest weight among the existing landmarks is selected and the value of the filter is multiplied.

As a result, at locations where landmarks do not exist in the vicinity, 1.0 point is added each time they are traveling; however, scores of 1.0 or more points are added at locations where landmarks are present in the surrounding area. Thus, the value of the location recognition score becomes high, and the influence of the landmark can be considered and quantitatively measured.

4.2.1 LANDMARKS WEIGHTING

As the impression of the location where the landmark exists is likely to remain, the presence or absence of the landmark greatly affects the construction speed of the cognitive map. However, the impression level given by the landmark varies depending on each landmark. It is necessary to quantitatively measure the landmark impression that is likely to be captured by a person.

Therefore, the landmarks are weighed according to their ease of recognition, and the rate of changing the value of the filter differed according to each landmark weight. The landmark ease of recognition is considered to differ depending on various factors such as type, size, and surrounding environment.

Fujii et al. [22] analyzed the randomly collected guide maps and investigated the elements that make up the guide map. The survey results assumed that the ease of recognition of landmarks varies depending on their attribute type (public facilities, restaurants, etc.), and the frequency of appearance is reported. As the landmarks that appear in guide maps, published in magazines, are generally recognizable landmarks, they are weighed in this study based on the landmark frequency of appearance.

According to Fujii et al., the landmarks with high occurrence frequency were stations, shops, restaurants, public facilities (post offices, libraries, parks, etc.), banks, and convenience stores. However, these landmarks are excluded because unlike walking, it is difficult to identify their external features during driving. The weight of the landmark is measured based on a value obtained by normalizing the landmark frequency of appearance that was reported by Fujii et al. to be in the range of 1.0 to 2.0. In other words, the value of the filter is 1.0 at the minimum (where no landmarks exists) and 2.0 at the maximum (when there is a landmark with maximum impression). Table 1 shows the landmarks adopted in the proposed method, their

frequency of occurrence, and the calculated landmark weights.

4.3 LOCATION RECOGNITION LEVEL MODULE

When assuming that the location recognition level is utilized by LBSs, it is also necessary to determine a location recognition level not only at the point where the user is present but also at the destination. Therefore, this proposed method can estimate the recognition level at every position.

The processing procedures of this module are summarized as follows:

1. Receive the location designation to estimate the location recognition level.
2. Locate the specified position and obtain the value recorded there.
3. Compare the acquired value with the threshold and output values of the three stages defined in Section 3 as the recognition level at the location.

A designated location was adopted. This is possible because the recognition level of each spot is recorded as a

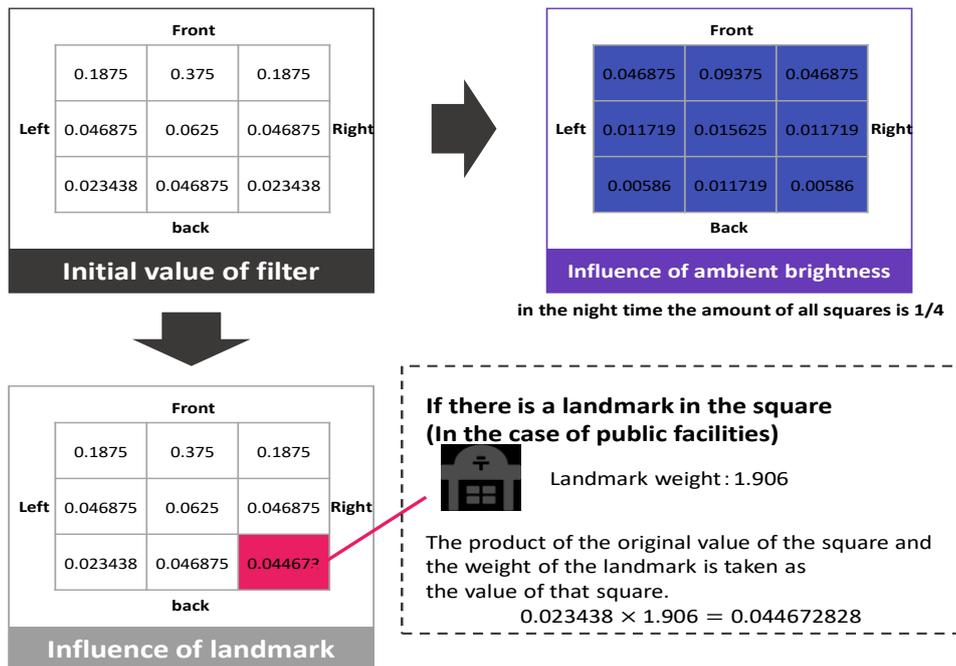
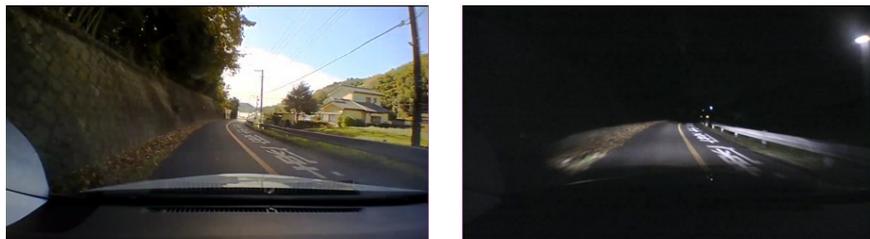


Figure. 4. Location Recognition Score Recording Module



Visibility at the moment when there is little light such as the street



Visibility at points where there are many lights such as streets

Figure. 5. Difference in visibility of day and night

score.

In this module, the score recorded at the designated location is compared with the threshold value, to estimate the recognition level at the specified location, which is obtained as an output.

The estimated location recognition level is measured in terms of the three stages defined in Section 3, and it is either “well recognizing,” “recognizing to a certain extent,” or “almost not recognizing.” Thresholds are set for each stage and used for estimating the recognition level. However, the appropriate thresholds vary depending on the users’ capabilities. The thresholds used in this paper are described in the next section.

The processing procedures of this module are summarized as follows:

1. Receive the location designation to estimate the location recognition level.
2. Locate the specified location and obtain the value recorded there.
3. Compare the acquired value with the threshold and output values of the three stages defined in Section 3 as the recognition level at the location.

4.4 IMPLEMENTATION

A system was implemented to estimate the location recognition level based on the proposed method. The evaluation described in the next section was conducted using the implemented system. The implementation environment is as follows:

- OS: Windows 10
- CPU: intel Core 5
- Programming language: Java (JDK 1.8)
- Save location recognition score format: XML

Acquisition of position information was performed using self-made application installed on NTT Docomo’s Optimus it L-05E. This application acquires position information every second using GPS and store latitude, longitude, and time stamp in CSV format.

To obtain landmark information, we used Place Library [23], a library of Google Maps API.

The implemented system runs on PCs. However, it can be easily migrated to applications running on Android terminal

Table 1 Frequency of appearance and weight of landmark

Attribute	Frequency of appearance	Weight
Station	0.117	1.898
A shop	0.13	2.000
Restaurant	0.018	1.614
Public facility	0.118	1.906
Bank	0.124	1.935
convenience store	0.049	1.362

because it uses Java language. As the location recognition level is assumed to be used in the LBS provided by smartphones, the location recognition level estimated by the proposed method can be adapted to applications that use location recognition level.

Considering the moving speed of the car and the viewing distance of the driver, the recording interval of the location recognition score (the interval when dividing the map into a lattice shape) was divided into 25 m × 25 m squares. There are several methods of calculating distances using latitude and longitude; however, Hubery’s formula, which can be easily calculated was selected[24].

5 EVALUATION AND CONSIDERATION

5.1 EVALUATION

To verify the method developed in this study, the estimated results were compared to experimental data that were accurately measured using traveling video. The conformity rate that compares the estimated results of the location recognition level and the accurately measure experimental data is used as an evaluation index. The outline of the verification process is shown in Fig. 6.

5.1.1 CREATING ACCURATE DATA

A travel video was used to create accurate data for two reasons, the experimental environment and the holding state of the cognitive map. When conducting experiments on actual vehicles, it allows to unify all the context conditions including the weather, brightness, and traffic conditions in the surroundings. Differences in experimental environments may affect the evaluation results. Hence, running images that can unify the experimental environment were used. A traveling image that can be presented as a visual image was used to allow remembering the location even though the state of the cognitive map at each of the divided points is at the stage of “understanding like a route map.”

Accurate experimental data were created from the results obtained by imposing multiple tasks on the participants who viewed the travel video. The task execution environment is shown in Fig. In order to evaluate whether or not the location recognition level with increased familiarity of the same route was correctly estimated, the travel video was viewed while executing the task once a day for 5 days in all. The tasks performed by the participants are as follows.

1. Extraction of stored landmarks
Participants enumerate the stored landmarks from among the viewed traveling images.
2. Storage confirmation of the scene

Scenes of the route where the travel video was viewed are randomly selected and presented as one image. (The presented image was taken as a daytime image regardless of whether the viewed travel video was during the day or night.) The participants report whether they remember the scenes presented. If the participant does not remember the presented scene, that location is reported as “almost not

recognized” in the experimental data. If the participant remembers the scene, he/she should decide the recognition level whether it is “recognized to some extent” or “well recognized” by the following task in which the image was presented 100 times each time.

If the context of the two scenes has been recognized correctly, the recognition level at that location is set as “well recognized” in the experimental data. If the context of the two scenes has not been recognized correctly, the location is reported as "recognized to a certain extent."

The video was recorded using a drive recorder. Both daytime and nighttime images were prepared to evaluate whether the captured image can cope with the difference in the surrounding environment. In addition, the video was recorded in two traveling directions, going and returning, in order to evaluate whether it was able to cope with the difference in the scenery in the reversed travel direction. The travel video was recorded in three routes shown in Fig. 5.3, and it was taken as the route of first seeing for the subjects. When the travel video was shot, the position

information was also acquired using the GPS. To acquire the location information, a self-made application was developed to record the location latitude, longitude, and time stamp in CSV format. After recording, the video was divided into 1 s snapshots and the position information corresponding to each snapshot was used.

5.1.2 EXPERIMENT

A total of 12 subjects participated in the experiment, 11 males and 1 female. All participants were in their twenties. Nine of the participants possessed an ordinary driving license and have road driving experience. In addition, the participants were students in the Informatics Department and the Faculty of Automotive Engineering. Approximately, half of the participants were using cars from the usual way. The participants were divided into 4 groups of 3 participants. The experiments involved the groups watching the video with different travel directions/surrounding environment (brightness).

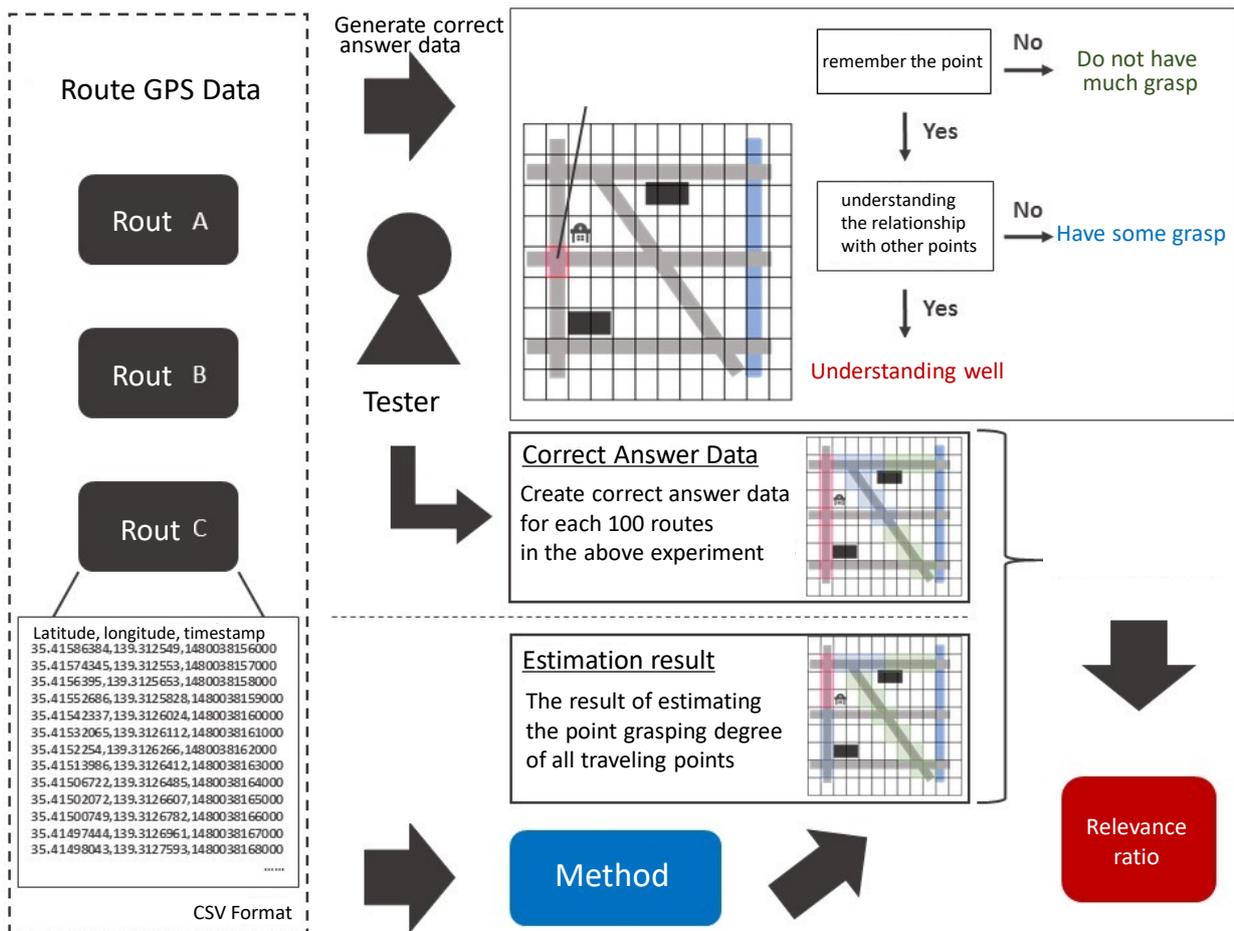


Figure. 6. Evaluation summary

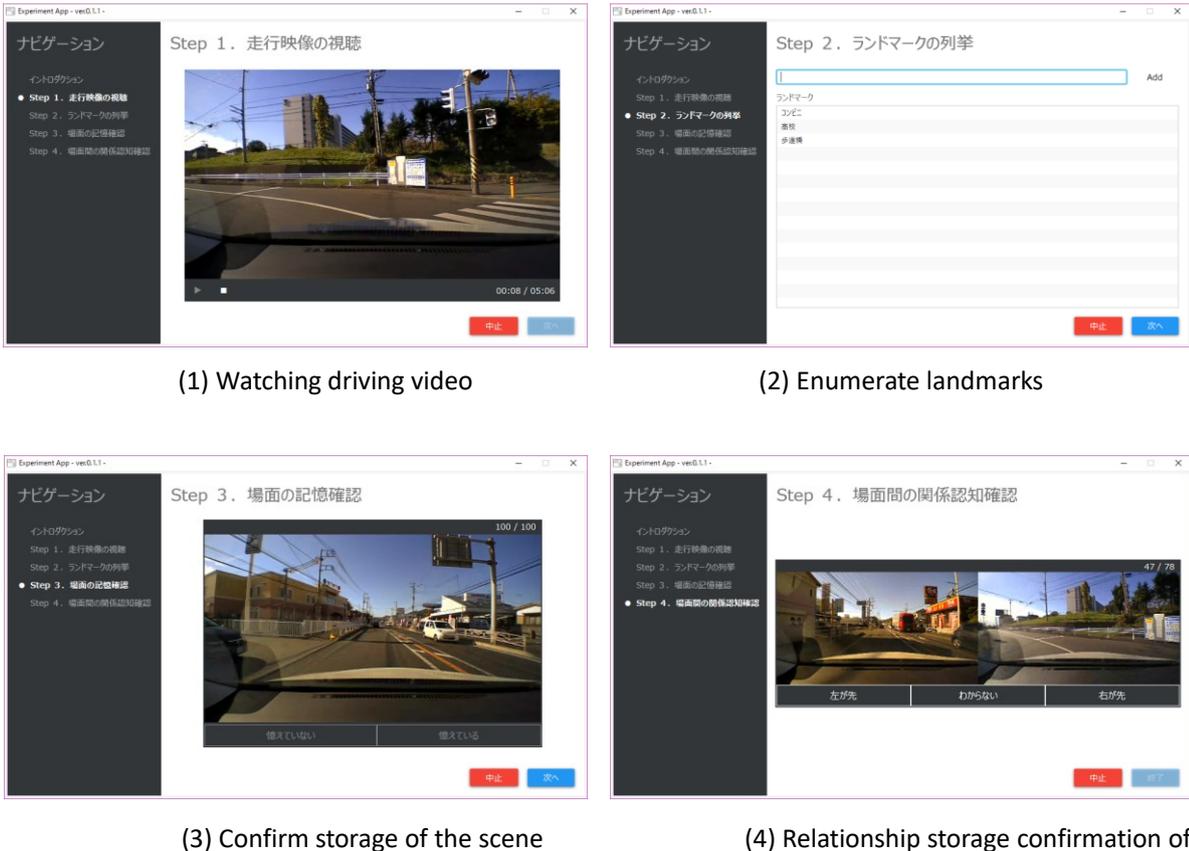


Figure. 7. Experiment environment

5.2 RESULTS AND DISCUSSION

5.2.1 INFLUENCE OF FAMILIARITY AND LANDMARK

The number of scenes that the participant remembers every time the viewing frequency is superimposed on any route in the memory confirmation task of the scene, which group A participants viewed, is increased after viewing each video. This trend is commonly observed among the participants in all groups. This result confirmed the effectiveness of the proposed approach to replace the concept of familiarity, which is reported to greatly affect cognitive mapping, with a form that can be quantitatively measured in terms of travel frequency.

In addition, it is confirmed that the number of scenes remembered in each route is different. At locations with high concentration of landmarks, such as the national highway, it is difficult to recognize each landmark individually while traveling. Thus, even if there are landmarks, their influence on the space recognition memory

becomes small. On the other hand, there are hardly any feature that can be considered as landmark on the route that goes.

5.2.2 COMPLIANCE RATIO OF PROPOSED METHOD

The average of all the relevance ratios that the proposed method obtained for each route and each viewing number is 49.7%. It is believed that the achieved accuracy by the developed method to estimate location recognition level is acceptable for information service applications. However, the accuracy should be further improved to be used for applications that affect safety, such as those used for supporting terminal operations.

6 CONCLUSION

In this study, a method was proposed to estimate the drivers' location recognition level using smartphones. The location recognition level is a state that shows how much the

user can recognize a specific location. The location recognition level is quantitatively measured based on the cognitive map building processes. Therefore, the study focused on the concept of "cognitive mapping." A cognitive map is a concept that refers to a mental image of the space that a person develops. The relative literature in the fields of cognitive psychology and applied psychology, which examined the factors that affect the process of space recognition, was reviewed. It has been reported that familiarity (experience of space) and landmarks greatly influence the process of building cognitive maps. Familiarity is the location information that can be acquired using GPS and the like. Meanwhile, the landmark information can be acquired via the Internet. The recognition level at each point is scored, and the state of the user's recognition level is quantified. The proposed method was used to quantitatively measure the factors that affect the location recognition level in a similar way.

ACKNOWLEDGEMENT

A Part of this work was supported by JSPS KAKENHI Grant number 16K00143

REFERENCE

- [1] http://www.kantei.go.jp/jp/singi/it2/senmon_bunka/detakatsuyokiban/dorokotsu_dai1/siryoushu3.pdf
- [2] Ministry of Internal Affairs and Communications: White Paper Information and Communication in Japan
- [3] S.Matsuyama,, Y. Tokunaga, R. Kiyohara," A Proposal and Evaluation of Method for recognizing User's Cognitive Levels of Location Focused on the Cognitive Map," IPSJ-ITS, 2016-ITS-66(12), 1-7 (2016-09-07) , 2188-8965
- [4] S. Matsuyama, Y. Tokunaga, R.Kiyohara, A Proposal and Application of Method for recognizing User's Cognitive Levels of Location Focused on the Cognitive Map, ITS symposium 2016(ITS Japan)
- [5] Seiji Matsuyama, Yuichi Tokunaga and Ryozi Kiyohara, Recognizing User Location Context for Car Navigation Devices IEEE International Symposium on Consumer Electronics 2016
- [6] Wei Han, Yan Yang, Guang-Bin Huang, and Cornelia Denk. Driver drowsiness detection based on novel eye openness recognition method and unsupervised feature learning. Systems, Man, and Cybernetics (SMC), 2015 IEEE International Conference on. IEEE,2015.
- [7] Guosheng Yang, Yingzi Lin, and Prabir Bhattacharya. A driver fatigue recognition model based on information fusion and dynamic bayesian network. Special Issue on Intelligent Distributed Information Systems, Vol. 180, pp. 1942-1954, 2010
- [8] Qiang Ji and Xiaojie Yang. Real-time eye, gaze, and face pose tracking for monitoring driver vigilance. Real-Time Imaging, Vol. 8, pp. 357-377, 2002.
- [9] Boon-Giin Lee and Wan-Young Chung. A smartphone-based driver safety monitoring system using data fusion. sensors, Vol. 12, pp. 17536-17552, 2012.
- [10] Edward C. Tolman. Cognitive maps in rats and men. The Psychological Review, Vol. 55, No. 4, pp. 189-208, 1948.
- [11] H.XU, S. MATSUSHITA, K. NISHIDE1, CHARACTERISTICS OF COGNITIVE MAPS : Simulation study on pedestrian path choice and spatial cognition part 2, Journal of Architecture and Planning (Transactions of AIJ) No. 545 p. 173-179
- [12] Kevin Lynch. The image of the city. The MIT Press, 1960.
- [13] F. Shemyakin. Orientation in space. Psychological Science in the USSR, Vol. 1, pp.186-225, 1962.
- [14] D. A. Appleyard. Styles and methods of structuring a city. Environment and Behavior, Vol. 2, pp. 100-116, 1970.
- [15] Alexander W. Siegel and Sheldon H. White. The development of spatial representations of large-scale environments. Advances in Child Development and Behavior, Vol. 10, pp.9-55, 1975.
- [16] N. Shingaki, "Why do People get lost?: Features of the Cognitive Process during a Second Visit,". Cognitive Studies, Vol. 5, No. 4, pp. 108-121, 1998.
- [17] K. Ihara, J SHoji. Cognitive map for Kids: development of cognitive map (1). edpsych1980, Vol. 22, pp. 250-251, 1980.
- [18] George A. Miller. The magical number seven, plus or minus two. The Psychological Review, Vol. 63, pp. 81-97, 1956.
- [19] T. Mori, M. Saburi, T. Saito, K. Yamamoto1," Some Limitations of Memory in Cognitive Map and a Bias in the Shape of Drawn Map," The Journal of the Institute of Television Engineers of Japan, Vol. 45, pp. 1583-1588, 1991.
- [20] H. Kubota, A. Kato, Y. Kubota. Drivers' Cognitive Maps and their Behavior. INFRASTRUCTURE PLANNING REVIEW, Vol. 9, pp. 61-68, 1991
- [21] K. Fujii, K. Sugiyama. Route Guide Map Generation System for Mobile Communication. IPSJ Journal, Vol. 41, No. 9, pp. 2394-2403, 2000.
- [22] Google: Place Library. <https://developers.google.com/maps/documentation/javascript/places/>, (accessed 2017/1).
- [23] Hubeny Karl. Zur entwicklung der gauss'schen mittelbreitenformeln. Österreichische Zeitschrift für Vermessungswesen, Vol. 42, pp. 8-17, 2012

VIS performance analysis in multi-lane T junction

Hiroki Ito^{*}, Yusuke Takatori^{**}

Kanagawa Institute of Technology, Japan

^{*}s1412062@cce.kanagawa-it.ac.jp, ^{**}takatori@ele.kanagawa-it.ac.jp

Abstract—In this study, we analyze the influence of the number of lanes on the information acquisition rate of a vehicle information sharing (VIS) system. VIS is a system in which vehicles with an on-board unit (OBU) for vehicle-to-vehicle (V2V) communication share their own information by broadcasting them to one another. Unless the OBU is connected to all vehicles, a mixed traffic of OBU- and non-OBU-equipped vehicles becomes an important traffic situation. In particular, because of communication error in a media access control layer, the number of lanes can affect the VIS performance in case the OBU penetration rate or traffic fluctuates. To analyze the VIS performance, we construct two traffic simulation scenarios, namely, one- and two-lane priority-road scenarios. Each scenario is constructed on the Scenargie 2.0 simulator. The simulation results show that the increase in the number of lanes significantly influences the fluctuation of VIS performance when the OBU penetration rate or traffic volume varies.

Keywords: ITS, vehicle information sharing (VIS), V2V communication

1 INTRODUCTION

Intelligent transport system (ITS) is a system that advances the movement of people and objects using information communication technology. Driving-assistance system that reduces accidents caused by vehicles is one ITS application expected to widely and rapidly spread. In particular, detecting obstacles using vehicular on-board sensors that cannot be seen from a vehicle is difficult. Meanwhile, when vehicles share their front-seat information (such as location, velocity, and other information) via vehicle-to-vehicle (V2V) communication, they can identify obstacle (other vehicle) information even they do not see each other. One of the applications that used such vehicle information sharing (VIS) system is the intersection collision warning system (ICWS) [1]. However, if vehicles with an on-board unit (OBU) for V2V communication (“OBU vehicle” for short) and a non-OBU vehicle are present, naturally, the non-OBU vehicle information cannot be shared. Therefore, a performance evaluation of VIS is important to penetrate such a system. In [1], the authors theoretically analyzed the performance of VIS in terms of OBU penetration rate. However, the theoretical analysis assumed an ideal communication environment and did not consider the influence of communication errors. Meanwhile, a standard of V2V communication, which is expected to spread in the future (IEEE802.11p [2]), considers carrier sense multiple access/collision avoidance (CSMA/CA) as its medium access control (MAC). Because the VIS transmission amount increases with OBU penetration, the frequency of packet collision in VIS may significantly depend on the OBU penetration rate. In previous research, a

microscopic traffic simulator that considered MAC was constructed, and the VIS performance for ICWS in terms of OBU penetration rate was analyzed using simulation [3]. Its result indicated that the VIS performance deteriorates if the OBU of VIS has a broad penetration.

However, these studies were performed in typical road environment. To apply them to practical use, we need to deal with a more complicated road environment. In particular, in MAC, as the number of road lanes increases, traffic volume fluctuation that accompanies the traffic volume increases. Therefore, analyzing the performance of VIS in terms of the number of lanes is important. In this study, we analyze the influence of the number of lanes on the information-acquisition performance of VIS by considering MAC.

This paper is organized as follows. Section 2 explains the VIS analysis by considering MAC. Section 3 constructs a simulation setup for VIS performance analysis in a multi-lane road environment. In Section 4, a simulation is performed to analyze the VIS performance. Finally, in Section 5, we conclude this paper.

2 VIS ANALYSIS CONSIDERING MAC

2.1 System overview

The VIS system is a system that shares vehicle information using V2V communication. In this research, WAVE [4] (IEEE802.11p + IEEE1609.X) is assumed as the environment for a dedicated short-range communication for vehicles. In WAVE, the vehicle with an OBU broadcasts a basic safety message (BSM) packet to surrounding vehicles in a short cycle (10 Hz). The BSM is defined in SAEJ 2735 “Dedicated Short Range Communications message set” [5]. It contains vehicular front-seat information. Information on the surrounding vehicles can be used to support safety driving. In WAVE, the physical, data-link, and application layers of an open systems interconnection reference model are standardized. Fluctuations in the number of nodes due to traffic-volume fluctuation and penetration rate of OBUs affect the MAC in the data-link layer. Therefore, considering MAC is important when the VIS performance is evaluated. Hence, we focus on the influence of MAC on VIS.

CSMA/CA is a MAC protocol for carrier transmission in the 802.11 networks. [6] The procedure of MAC by CSMA/CA is explained as follows:

- Carrier sensing: Prior to transmission, a node (vehicle) first listens to the shared media (such as listening for wireless signals in a wireless network) to determine whether another node is transmitting or not.

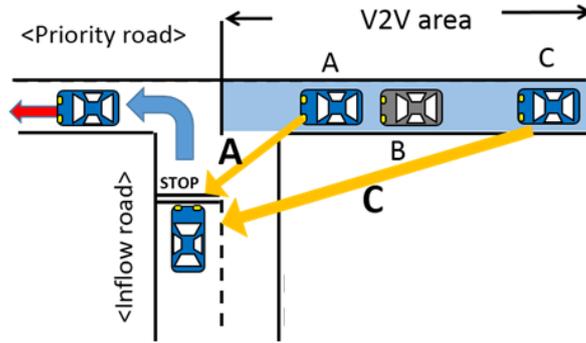


Fig. 1: VIS for ICWS (Operating environment)

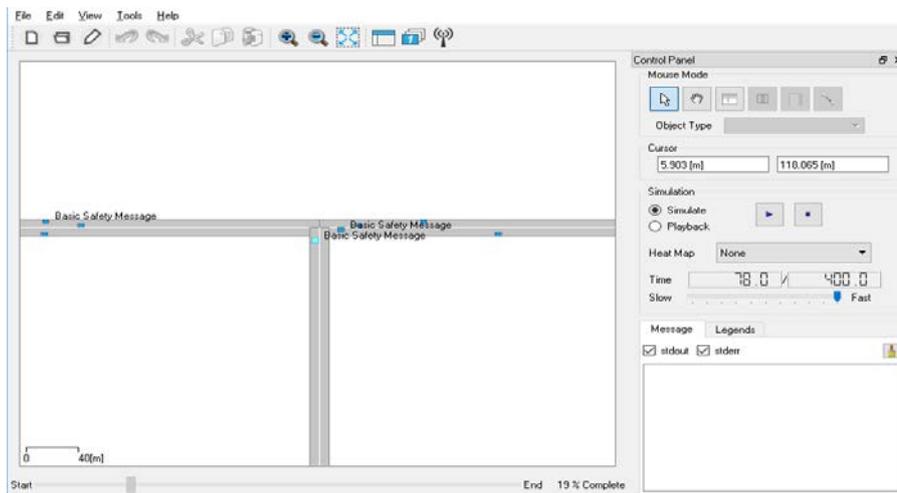


Fig. 2: Scenargie GUI

- Collision avoidance: If another node is sensed, the node waits for a period of time (usually random) for the other node to stop transmitting before listening again for a free communication channel.
- Transmission: If the medium is identified as clear, the waiting node sends the frame in its entirety.

CSMA/CA achieves high throughput. However, if the number of nodes or the packet size increases, its performance decreases with the increase in the packet collision probability.

2.2 Performance analysis of VIS

In the previous research, the performance of the VIS system for ICWS on a T-shaped road was assumed as a target application. ICWS discourages a driver in a vehicle that approaches a blind intersection from entering the intersection if another vehicle approaches the same intersection from another direction. The operating environment of the application is shown in Fig. 1.

The intersection combines a priority road and an inflow road. The road that horizontally extends is called the priority road, and vehicles traveling on it are called priority vehicles. In the priority road, priority vehicles travel on the left-hand side of the road (drivers follow Japanese traffic rules). Furthermore, the road that intersects at vertical angles is called the inflow road, and vehicles traveling on it are called inflow vehicles.

Inflow vehicles should not tie up with priority vehicles that approach the intersection.

When the line-of-sight from the inflow road to the priority road is interrupted by obstacles (pedestrians, hedges, walls, buildings, etc.), the driver of the inflow vehicle may pass over the priority vehicles and create a dangerous situation that may result in a crossing collision. To avoid this danger, the ICWS assists the inflow vehicle in safely entering the intersection by providing driving information (position, velocity, etc.) of the priority vehicles approaching the intersection. In this system, for the priority vehicle to repeatedly broadcast its individual driving information at short intervals, the OBU of the system must have two fundamental functions; a real-time V2V communication function and a real-time positioning function. By using these functions, the inflow vehicle OBU can acquire the driving information of the other priority vehicle in real time.

The mean of the entire vehicle information acquiring probability (MEVIAP) is considered as an evaluation index for analyzing the VIS performance [1]. MEVIAP represents a probability that an inflow vehicle receives information on each other vehicle every 0.1 s and can receive all vehicle information existing in the V2V communication area during the measurement time of the entire vehicle information

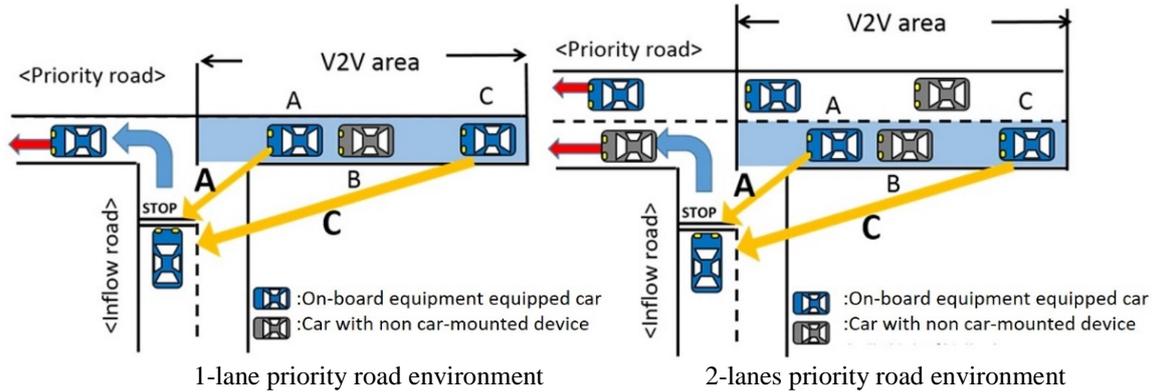


Fig 3: Road situations

Table 1: Simulation conditions

Road Configuration	T-junction with 1-lane priority road T-junction with 2-lanes priority road
Distribution of Headway distance	Exponential distribution
V2V area	200m from the center of the intersection
Penetration rate	0 - 100[%]
transmission interval	100[ms]
Payload size	128[byte]
Vehicle speed	40[km/h]
Average traffic flow	150, 300[vehicle/h]

acquisition probability. Higher MEVIAP reduces the number of approaching vehicles that are not recognized by the inflow vehicles and considers them as among the vehicles approaching the intersection.

2.3 Scenargie simulator

For the VIS performance analysis, we need to simultaneously simulate communication and car traffic.

In this research, we performed the analysis using MAC (CSMA/CA) of IEEE 802.11p and Scenargie 2.0 [7] in which a traffic-flow model is implemented. The Scenargie graphical user interface is shown in Fig. 2.

3 CONSTRUCTION OF SIMULATION ENVIRONMENT

To analyze the influence of the number of lanes, two road situations shown in Fig. 3 were constructed on the simulator.

Assuming that the vehicle in the inflow road is waiting to turn left, the position will not change. The vehicle in the priority road travels while transmitting its own information to the standby inflow vehicle.

To construct this simulator, the V2V communication module [based on IEEE 802.11p and IEEE 1609 (WAVE)] provided by Scenargie 2.0 was then implemented in the simulator.

To analyze the performance of the mixture of OBU and non-OBU vehicles, two types of vehicle models were prepared. One was an OBU vehicle model with an OBU

composed of V2V and forward vehicle-detection functions. The other was a non-OBU vehicle without an OBU.

In this simulator, communication is simulated in a sub-nanosecond order, and the communication result is output every 0.1 s, which is the update timing of vehicular dynamics. Both communication and vehicular dynamic results are output to a trace file. Then, the result is used to analyze the VIS performance.

In this analysis, we focused on the priority lane that directly connected the inflow road. In other words, the right lane of the two-lane priority-road was not an analysis target. The simulation conditions are listed in Table 1. The precise definition varies depending on the application, but it is most commonly measured as the distance from the tip of one vehicle to the tip of the next vehicle behind it, expressed as the time it takes for the trailing vehicle to cover that distance.

4 SIMULATION RESULTS

The MEVIAPs of the two types of roads are analyzed in terms of the penetration rate. In each road environment, two types of traffic flow (150 and 300 vehicles per hour) are analyzed. The results are shown in Fig. 4. Because our previous research [8] indicated a MEVIAP of more than 60% of the driver confidence, we evaluate the VIS performance at a penetration rate of 60% or more. Moreover we also evaluate the MEVIAP at 0% penetration for comparison. Fig. 4 shows that the MEVIAP increases with the OBU penetration. In the case where the OBU penetration increases to more than 60%, the MEVIAP of the two-lane priority road is lower than that of the one-lane priority road. Moreover, in both road

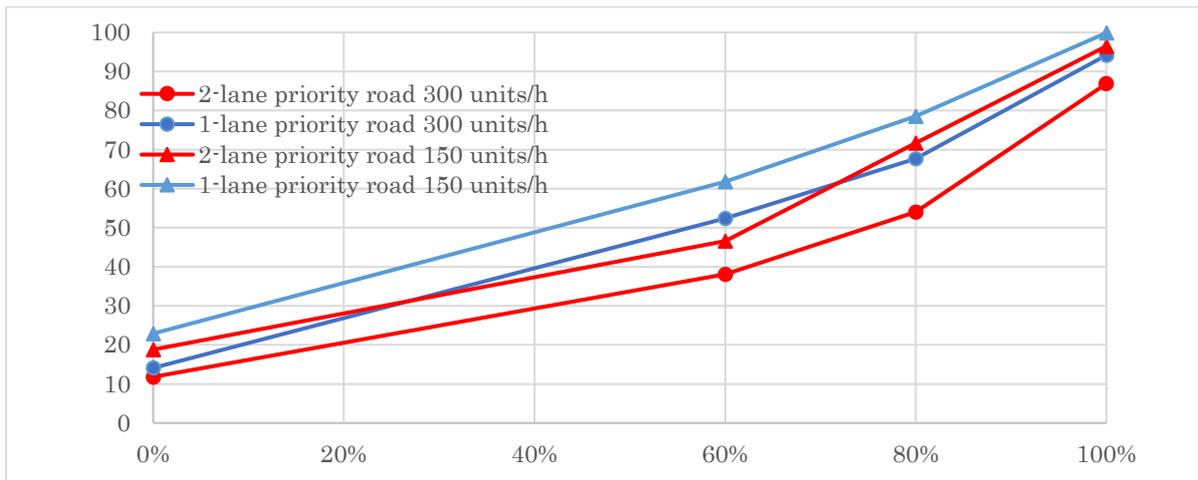


Figure 4: Result of simulation

conditions, the MEVIAP decreases with the increase in the traffic flow. In particular, the deterioration ratio of the MEVIAP of the two-lane priority road caused by increasing traffic flow is larger than that of the one-lane priority road. These results indicate that the influence of OBU penetration on the MEVIAP depends on the number of priority lanes.

Moreover, the deterioration of the MEVIAP caused by the increase in the traffic flow depends on the number of lanes in the priority road.

5 CONCLUSION

In this research, we analyzed the influence of the increase in the number of lanes on the VIS performance using a simulator. The results of the performance analysis show that the number of priority lanes significantly affects the acquisition of vehicle information.

Because we assumed that the OBU vehicles broadcast only its individual front-seat information in this study, in future research, we will analyze the influence of the number of lanes in a priority road in the case where OBU vehicles broadcast not only the information of its individual information but also that of surrounding vehicles detected by its obstacle detection sensors.

REFERENCES

- [1] Yusuke Takatori, Hideya Takeo, "Analysis of Vehicle Information Sharing Performance of an Intersection Collision Warning System", IEICE TRANS. FUNDAMENTALS, VOL.E100-A, NO.2 FEBRUARY 2017.
- [2] IEEE 802.11p-2010," IEEE Standard for Information technology-- Local and metropolitan area networks-- Specific requirements-- Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 6: Wireless Access in Vehicular Environments"
- [3] Y.Takatori, S.Onodera, T.Sugiyama, R.Kiyohara,

" Analysis of Vehicle Information Sharing System by Microscopic Traffic Flow Simulation", Proc . of International Workshop on Informatics , pp83-87 , 2016.

- [4] Roberto A. Uzcátegui," Wave: A tutorial", IEEE Communications Magazine, Volume 47, Issue 5, pp.126-133, (2009),
- [5] SAE J2735," Dedicated Short Range Communications (DSRC) Message Set Dictionary", (2009).
- [6] L. Kleinrock and F.A. Tobagi, "Packet switching in radio channels: Part 1: CSMA modes and their throughput-delay characteristics," IEEE Trans. Commun., vol.COM-23, no.12, pp.1400-1416, 1975
- [7] Scenargie: Space-Time Engineering, LLC , <http://www.spacetime-eng.com/>
- [8] Yusuke Takatori, Hideya Takeo," Evaluation of a driver acceptance of a vehicle information sharing on the penetration rate", Proc.of. 22nd ITS World Congress, Paper number ITS-2154, Bordeaux, France, 5-9 October 2015

Evaluation of Embedded Platform for Autonomous Drive Using Linux and GPU

Shota Mizuno[†], Shinya Honda[‡], Ryo Kurachi[‡], Daisuke Hayakawa* and Naoya Chujo**

[†]Graduate School, Aichi Institute of Technology, Japan

[‡]Graduate School, Nagoya University, Japan

* Panasonic Industrial Devices SUNX Co., Ltd., Japan

** Aichi Institute of Technology, Japan

{b16727bb,ny-chujo}@aitech.ac.jp

honda@ertl.jp

kurachi@nces.is.nagoya-u.ac.jp

hayakawa.daisuke@jp.panasonic.com

Abstract - In recent years, autonomous driving systems and advanced driver assistance system (ADAS) have been developed to achieve an improved level of safety. In response to the increasing complexity of the software in these systems, automakers and suppliers have developed a standardized operating system called Automotive Grade Linux (AGL) for vehicle infotainment systems. In addition, the GPU is a promising device that can be used to meet the demand for high computing power required for advanced image processing and deep learning. However, to the best of the authors' knowledge, the real-time performance for ADAS using Linux and a GPU has not yet been reported.

In this study, two platforms for an ADAS using Linux and a GPU were examined. An ADAS containing an image processing board with a GPU and using the Linux operating system was developed and implemented using a miniature car. The present experimental results demonstrate that this system can successfully achieve real-time image processing. In evaluations of the ADAS using a GPU, it was found that bottleneck of GPU processing was the image capture process.

Keywords: Autonomous driving, Advanced driver assistance systems, Linux, Real-time property, GPU

1 INTRODUCTION

In recent years, autonomous driving systems and advanced driver assistance systems (ADAS) have been actively developed to achieve an improved level of safety [1][2]. In each autonomous driving system and ADAS, a large number of sensors, such as stereo cameras, light detection and ranging (LIDAR) sensors, and radar, are used to recognize various driving environments. It is also necessary to simultaneously process a large amount of information from these sensors in real time because vehicles must be capable of traveling on highways.

For these reasons, autonomous driving systems and ADAS require a standard environment for software development and high-speed hardware. Regarding the standardization of software, automakers and suppliers have developed a standardized operating system called Automotive Grade Linux (AGL) [3], which is a powerful technology for today's complex systems. Linux OS contains abundant device drivers, APIs, and libraries. In addition, the latest multicore CPUs and their mul-

titasking are supported. There have been attempts to improve the real-time nature of Linux using RTLinux [4], which is a Linux kernel with a dedicated patch applied. With regard to hardware, using a large GPU is a promising technology candidate for CPU. The development of the semiconductor technology has dramatically accelerated the advancement of GPUs with a massively parallel architecture. Additionally, GPUs seem to be essential for the improvement of the recognition performance of ADAS using deep learning. However, to the best of the authors' knowledge, the evaluation of the real-time performance of Linux for autonomous driving systems has not yet been reported. There have also been no reports on the effect of acceleration of the recognition of the road environment using GPUs.

In this study, two platforms for autonomous driving were developed using Linux and a GPU, and the factors affecting the real-time performance of the platforms were examined. In addition, the real-time processing capabilities of the embedded GPU were verified. The aim of these evaluations was to obtain guidelines for implementing software for the realization of a real-time system using Linux. With respect to hardware, additional aims were to evaluate the potential of the GPU and better understand the bottleneck of image processing for autonomous driving and how to maximize the parallelism of the GPU.

The remainder of this paper is organized as follows. Section 2 describes two platforms based on the architecture for autonomous driving systems and the elements comprising these platforms. Section 3 discusses experiments that were conducted to verify the platform using Linux and a real-time OS (RTOS) and the real-time property using it. In Section 4, experiments on platforms based on Linux, the RTOS, and a GPU are presented. The experimental considerations are discussed in Section 5, and Section 6 concludes the paper.

2 PLATFORM FOR AUTONOMOUS DRIVING

Autonomous driving systems and ADAS contain a number of different sensors. On-board cameras are commonly used and are indispensable for recognizing lanes, pedestrians, and signs. Subsection 1 introduces the architecture for autonomous driving, and subsection 2 describes two platforms for evaluation.

2.1 Minimum architecture for autonomous driving

Fig. 1 shows the minimum architecture for an autonomous driving system, which consists of an on-board camera, an image processing system, and a vehicle control system.

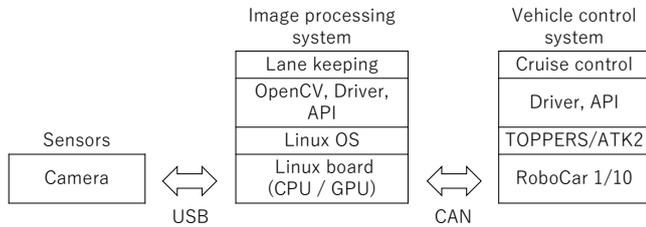


Figure 1: Minimum architecture for autonomous driving system

2.2 Two platforms for evaluation

On the basis of this architecture, two platforms were considered in this study. Both platforms A and B use Linux and an RTOS, but only Platform B is equipped with a GPU. In Platform A, relatively lightweight processing is performed, whereas in Platform B, advanced image processing requiring substantial power is performed. The features of two platforms are given in Table 1.

Platform A consists of a camera, an RTOS and a Linux OS. Raspberry Pi 2 Model B was adopted as the image processing system. This system consists of a small CPU board that runs on Linux-based OS and has a quad-core ARM Cortex-A7 processor and 1 GB of DDR-SDRAM. RoboCar [5], which was used for the vehicle control system, is a miniature version of an actual car. It has a V850 CPU board with a TOPPERS/ATK2 OS [6], which is an RTOS for automotive systems. Raspberry Pi and Robocar are connected via a controller area network (CAN). A universal serial bus (USB) camera (iBUFFALO BSW20KM11) with a resolution of 640×480 and a maximum frame rate of 30 fps is connected to the image processing system.

In platform B, Jetson TX1 is used as the image processing system. This system has 256 CUDA cores and a quad-core ARM Cortex-A57 processor with a 2 MB L2 cache and a 4 GB LPDDR4 memory. In this setup, Jetson TX1 and RoboCar are connected via a CAN. A USB3.0 stereo camera is connected to Jetson TX1. The considered stereo image resolutions were wide video graphics array (WVGA) (672×376), HD720 (1280×720), and HD1080 (1920×1080).

3 REAL-TIME PERFORMANCE WITH LINUX

The real-time performance of Linux was evaluated using an autonomous driving platform.

3.1 Lane keeping system

Lane keeping [7] is a basic driving support system, and its configuration is simple. A lane keeping system consisting of white line recognition and steering control was implemented in Platform A to verify the real-time performance of Linux. Fig. 2 shows a flow diagram of the lane keeping system.

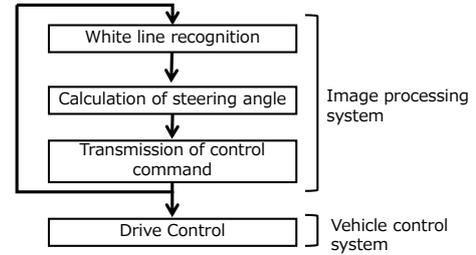


Figure 2: Flow diagram of the lane keeping system

A camera is used to recognize the white lines. According to the recognition result, the image processing system calculates the steering angle, and the result is sent to the vehicle control system.

3.2 Processor affinity for CPU selection

Processor affinity is a property used to specify which CPU executes each process. In Linux, the OS normally automatically assigns an execution processor based on this property. If the process can be executed by another processor, the execution processor may sometimes be reassigned. When the execution processor is reassigned, the process is delayed to allow the cache to be copied.

In this study, the affinity was set to execute the lane keeping program on core 2 of Raspberry Pi using the taskset command. The execution time with and without the specification of the execution processor was measured for 10000 frames. The results are shown in Fig. 3.

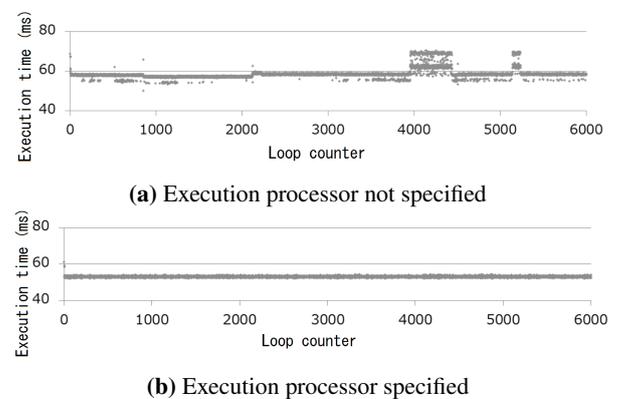
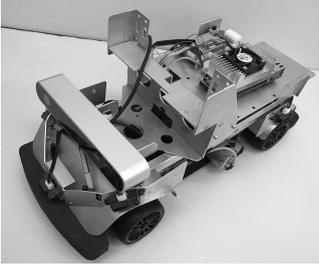


Figure 3: Execution time with and without the specification of the execution processor

As shown in Fig. 3a, lane keeping was typically executed

Table 1: Features of the two platforms

	Platform A	Platform B
		
Camera	iBUFFALO BSW20KM11	ZED Stereo Camera
Image processing system	Raspberry Pi 2 Model B CPU: 4× ARM Cortex-A7 GPU: —	NVIDIA Jetson TX1 CPU: 4× ARM Cortex-A57 GPU: 256× CUDA core
Vehicle control system	ZMP Inc. RoboCar 1/10	

within 60 ms when the execution processor was not specified. However, the execution time sometimes increased to approximately 70 ms in this case. This is thought to have been caused by the reassignment of the execution processor. According to Fig. 3b, this increase is suppressed when the execution processor is specified. The average execution time and the worst-case execution time (WCET) are given in Table 2.

Table 2: Average execution time and WCET with and without the specification of the execution processor

Execution processor	Not specified	Specified
Average	59.4 ms	53.1 ms
WCET	78.1 ms	61.3 ms

The average execution time and the WCET were lower when the execution processor was specified than when it was not specified, demonstrating that the WCET can be improved by the specification of the execution processor.

In a normal Linux OS, a timer interrupt occurs at regular intervals. When a timer interrupt occurs, the scheduler examines the process state, and the processes to be executed are selected in descending order of priority. Because the priority of a process that is using a CPU for a long time decreases with time, the timer interrupt enables the execution of other waiting processes. With this mechanism, Linux maintains the even execution of processes.

When lane keeping is executed without options, the priority of this process is the same as that of other processes. Therefore, the execution may be postponed by the scheduler in some cases. In addition, if the execution time of one cycle of the process is long, the scheduler is activated during the execution, and other higher-priority processes are executed. Then, the original process is delayed until it becomes executable again. It is thought that this is the reason why the execution cycle is delayed in a normal Linux OS.

3.3 Real-time process and RTLinux

RTLinux has been developed in an effort to improve the real-time performance of Linux. RTLinux with the RT-Preempt

patch applied was examined in this study. The RT-Preempt patch [8] is a patch provided by the Linux community to improve the real-time performance of the Linux kernel. Various types of patches exist for this purpose, including Xenomai and real-time application interface (RTAI) [9].

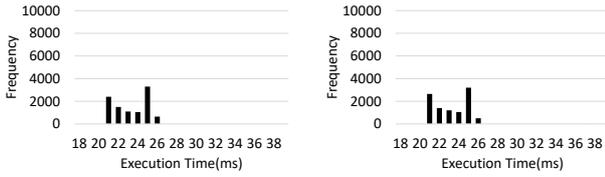
The lane keeping program described previously was implemented on the Platform A, and the execution time was measured. The cases with the normal and real-time processes were compared. In the case of the real-time process, the execution of the processor is given higher priority.

In RTLinux, the time required for the camera to capture the image was abnormally slow. This is likely because the device driver was updated when the RTLinux patch was applied. Therefore, the time required for image capturing was not considered here. Experiments were conducted for four cases with each possible combination of real-time or normal kernels and real-time or normal processes. The conditions for the four cases are given in Table 3.

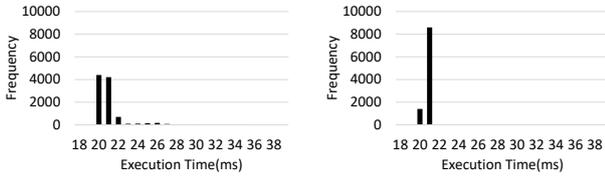
Table 3: Kernel and process conditions for each considered case

Case	a	b	c	d
Kernel	Normal	Normal	Real-time	Real-time
Process	Normal	Real-time	Normal	Real-time

Histograms of the lane keeping execution time obtained for these cases are shown in Fig. 4, and the average execution times and WCETs for the four cases are given in Table 4. In both Fig. 4 and Table 4, the time required to capture the image is excluded.



(a) Linux, normal process (b) Linux, real-time process



(c) RTLinux, normal process (d) RTLinux, real-time process

Figure 4: Histograms of the lane keeping execution time

Table 4: Average execution times and WCETs for lane keeping

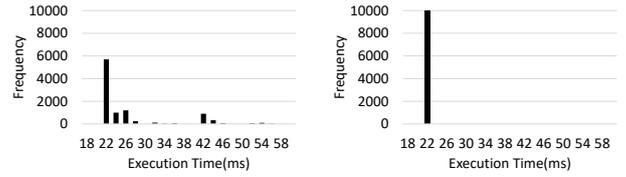
Case	a	b	c	d
Average	26.8 ms	20.5 ms	20.5 ms	20.3 ms
WCET	40.5 ms	21.5 ms	52.2 ms	21.7ms

When the program was executed as a normal process on RTLinux, the execution time was unstable in comparison with the normal process executed on the normal Linux kernel. However, when it was executed as a real-time process on RTLinux, the execution time was stable, and the delay was greatly suppressed. This is because the delay for dispatching the process is decreased by RTLinux.

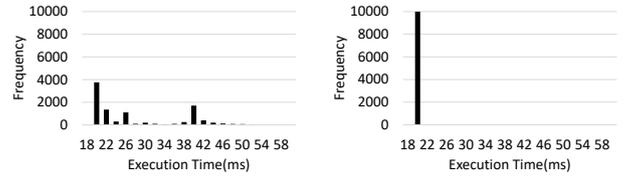
Applying the RT-Preempt patch reduced the time to start the scheduler. The real-time performance of the lane keeping program deteriorated in the kernel with the RT-Preempt patch applied without options. When the process is executed as a normal process, the execution state can be considered to switch more frequently. When the process is executed as a real-time process on RTLinux, it is preferentially executed by one of the cores, even if the execution core is not specified. The real-time performance was also improved in normal Linux by executing it as a real-time process.

3.4 Real-time process and RT Linux with CPU load

In the experiments discussed so far, only a single user process was executed. However, in an actual system, multiple processes are executed. Therefore, the delay in executing processes was examined while applying a load with background processing. For the same four cases as given in Table 3, a CPU load was applied using the stress command. Processing was performed for 10,000 periods. Histograms of the lane keeping execution time are shown in Fig. 5, and the average execution times and WCETs are given in Table 5.



(a) Linux, normal process (b) Linux, real-time process



(c) RTLinux, normal process (d) RTLinux, real-time process

Figure 5: Histograms of the lane keeping execution time with a CPU load

Table 5: Average execution times and WCETs of lane keeping with CPU load

Case	a	b	c	d
Average	25.9 ms	20.4 ms	27.4 ms	19.1 ms
WCET	67.4 ms	21.4 ms	57.3 ms	20.6 ms

According to Fig. 5a and c, the fluctuation of the execution cycle was larger in both kernels when the process was executed as a normal process. As shown in Table 5, the average execution time of the lane keeping program executed as a normal process in the normal Linux kernel (Case a) was 25.9 ms, and the WCET was 67.4 ms. In this case, the variation in the execution cycle was not evenly distributed but was concentrated in two ranges: 22–26 and 42–44 ms. In contrast, when executed as a real-time process in the normal Linux kernel (Case b), its average execution time and the WCET were 20.4 and 21.4 ms, respectively. Thus, executing the program as a real-time process stabilized the execution time.

When a CPU load was applied and the program was executed as a normal process, the execution cycle greatly fluctuated in both kernels. However, when it was run as a real-time process with a, n applied CPU load, the execution time was stabilized in both kernels. The execution time was stabilized by assigning the process a high priority, however, there are multiple high-priority processes in an actual system, it it a matter which process is prioritized.

4 REAL-TIME PERFORMANCE USING GPU WITH LINUX

This section discusses the evaluation of the real-time performance with a platform using a Linux board equipped with a GPU.

4.1 Lane keeping system

The lane keeping program was implemented on Jetson TX1 instead of Raspberry Pi in the image processing part; however, lane keeping is too lightweight to use a GPU. The image capturing time was excluded in the calculation of these results. The execution time of the lane keeping program was 5.4 ms, and the time was increased to 8.1 ms when using a GPU. Transfer between CPU and GPU occupies 1.7 ms. Because lane keeping is a lightweight process and the overhead of the data transfer is larger than the execution time reduction achieved by using the GPU.

4.2 Collision avoidance system

A collision avoidance system was implemented and evaluated by stereo image processing using Jetson TX1. Collision avoidance systems are being adopted by many manufacturers. Possible sensors used in such systems include millimeter-wave and infrared radar sensors and single and stereo cameras. A stereo camera was adopted in this study because it is the sensor that is best able to can detect the lane and identify the obstacles.

A flow diagram of the system is shown in Fig. 6.

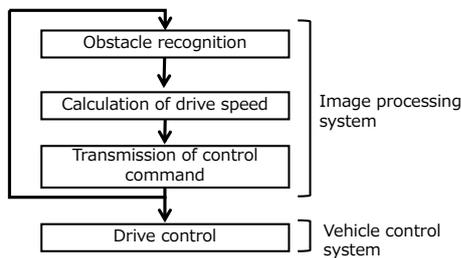


Figure 6: Flow diagram of the collision avoidance system

In this system, obstacle recognition and the calculation of drive speed are executed on the GPU, and the drive speed is then transmitted to the vehicle control system via a CAN. When an obstacle is recognized, the RoboCar is stopped.

4.3 Dependence on image resolution

Simple pedestrian detection has already been put to practical use, high-resolution images will be necessary to improve detection distance and estimation of pedestrian movement direction. When 8-mega-pixel images are used, it is also possible to detect pedestrians at distances of up to 200 m [10].

Several obstacle recognition experiments were conducted using stereo images at three resolutions: stereo of WVGA, HD720, and HD1080. Fig. 7 shows the processing speeds for each of the resolutions.

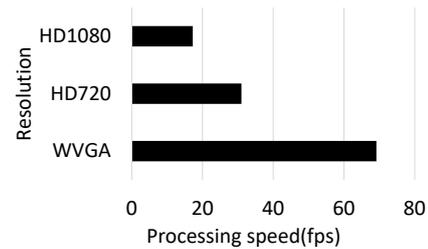


Figure 7: Comparison of processing speeds at different resolutions

The processing speeds for WVGA, HD720, and HD1080 were 68, 31, and 17 fps, respectively. Table 6 gives the processing speeds for the three resolutions relative to that of the WVGA case.

Table 6: Image sizes and processing speeds of collision avoidance on Jetson TX1 at different image resolutions relative to that of the WVGA case

Resolution	Image size	Processing speed
WVGA	1×	1×
HD720	3.7×	2.2×
HD1080	8.3×	4.0×

Although the image sizes of the cases with resolutions of HD720 and HD1080 were respectively 3.7 and 8.3 times that of the WVGA case, the processing speeds were only 2.2 and 4.0 times that of the WVGA case, respectively. This shows that the parallelization in the GPU is utilized. However, most of this execution time is spent on capturing the camera images. Fig. 8 shows the breakdown of the execution time of the collision avoidance system. The execution time in the GPU was measured synchronously.

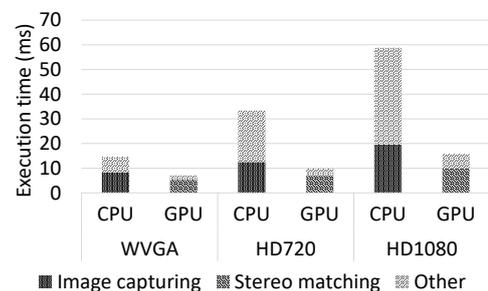


Figure 8: Execution time of the collision avoidance system

In the case of WVGA, the CPU execution time for one frame is approximately 15 ms, and approximately 10 ms of that time is spent on image capturing. The GPU was found to complete processing in 8 ms. At a higher resolution, the GPU execution time did not increase much. Thus, processing in the

CPU, particularly with regard to the capturing of the camera images, is a performance bottleneck.

4.4 Processor affinity for the process using a GPU

Experiments were also conducted to determine the affinity for Platform B. The recognition process was fixed to core 2 of Jetson TX1. The resulting execution time histograms are shown in Fig. 9. As in Section 3, the execution times of the collision avoidance program were compared with the program run as a normal process and as a real-time process.

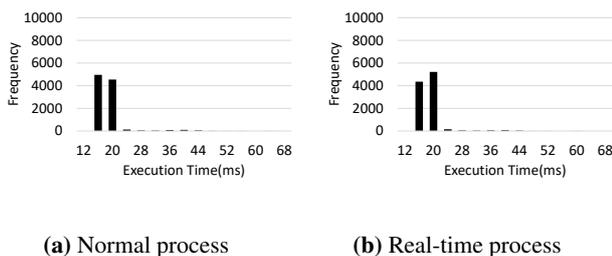


Figure 9: Histogram of the collision avoidance execution time

The execution times were concentrated near 16 and 20 ms for both the normal and real-time processes. In addition, even when the collision avoidance program was executed as a real-time process, the execution time was increased to nearly 60 ms in approximately 0.02% of the cases. This is because the program used in the experiment consisted of several sub-processes and sometimes other subprocesses were prioritized, causing the execution to be postponed.

5 CONSIDERATIONS

The experimental results obtained in the present study show that the real-time performance of software programs implemented in Linux can be improved by specifying the execution core without optimizing the programs. Executing programs as real-time processes is also effective in both the normal Linux and RTLinux kernels. Furthermore, by applying the RT patch to the Linux kernel, it is possible to stabilize the execution time. However, in an actual system, there are likely to be multiple high-priority processes. Therefore, it is necessary to design the priorities of processes and system calls and verify the system performance with sufficient test vectors.

With regard to hardware for autonomous driving using a GPU, in the image processing system used in these experiments, the operation executed on the GPU side was sufficiently fast, and the real-time property was considered to be satisfied. According to the development of GPU technology, the improvement of the image resolution and frame rate for autonomous driving are expected to be improved. However, the transfer time between the CPU and GPU and the bottleneck in the image capture process remains. Without an improved high-speed transfer channel and a high-speed interface with the camera, it is difficult to take advantage of the potential of using a GPU.

6 CONCLUSION

In this study, two platforms for autonomous driving were developed using Linux and a GPU, and autonomous driving applications were evaluated to investigate the real-time performance of the system.

In software using Linux, the real-time performance can be improved without optimizing the program by specifying the execution core of the program and executing the process as a real-time process. When the CPU is loaded with other processes, the process can be executed stably if it is executed as a real-time process.

With regard to hardware using a GPU, the present experimental results demonstrate that the evaluated collision avoidance system can achieve the real-time property for the recognition of the driving environment. The processing on the GPU side was executed at a sufficiently high speed. However, the execution time was not stable; it was found that the performance bottleneck of the system was the capturing of camera images.

Future work will include considering a more complex system to verify the performance of a practical autonomous driving system.

REFERENCES

- [1] Urmson, C., Anhalt, J., Bagnell, D., Baker, C., Bitner, R., Clark, M. N., ... & Gittleman, M., Autonomous driving in urban environments: Boss and the urban challenge, *Journal of Field Robotics* 25(8), pp.425-466(2008).
- [2] Lindgren, Anders, and Fang Chen. State of the art analysis: An overview of advanced driver assistance systems (adas) and possible human factors issues, *Human factors and economics aspects on safety*, pp.38-50(2006).
- [3] Linux Foundation, AUTOMOTIVE GRADE LINUX, accessed May 27(2017). <https://www.automotivelinux.org/>
- [4] Linux Foundation, Real-Time Linux, accessed May 27(2017). <https://wiki.linuxfoundation.org/realtime/start>
- [5] ZMP Inc, 1/10 scale RoboCar 1/10, accessed May 27(2017). <http://www.zmp.co.jp/products/robocar-110>
- [6] TOPPERS Project, TOPPES/ATK2, accessed May 27(2017). <https://www.toppers.jp/atk2.html>
- [7] Robert Bosch, BOSCH Automotive Handbook, (2011)
- [8] Arthur, Siro, Carsten Emde, Nicholas Mc Guire, Assessment of the realtime preemption patches (RT-Preempt) and their impact on the general purpose performance of the system, *Proceedings of the 9th Real-Time Linux Workshop*(2007).
- [9] Barbalace, Antonio, et al. Performance comparison of VxWorks, Linux, RTAI, and Xenomai in a hard real-time application, *IEEE Transactions on Nuclear Science* 55.1, pp.435-439(2008).
- [10] ADAS Front Camera: Demystifying Resolution and Frame-Rate, accessed May 27(2017). http://www.eetimes.com/author.asp?section_id=36&doc_id=1329109

Keynote Speech 1:
Prof. Ryoichi Sasaki
(Tokyo Denki University)



IWIN 2017
Keynote Presentation

SEP. 4 , 2017

Digital Forensics in Japan



Professor, Tokyo Denki University
Ryoichi Sasaki
sasaki@im.dendai.ac.jp



1

My Profile (1)

Dr. Sasaki received his B.S. Degree in health science and Ph.D Degree in system engineering from the University of Tokyo in 1971 and 1981, respectively.

From April of 1971 to March of 2001, he was engaged in the research and research management related to systems safety, network management and information security at Systems Development Laboratory of Hitachi Ltd.

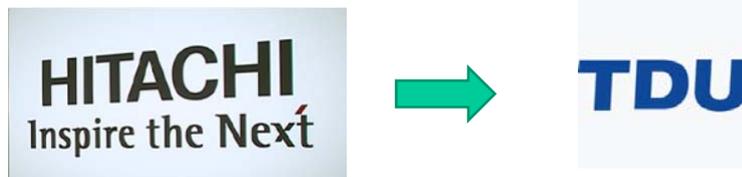


2

My Profile (2)

Dr. Sasaki started the study of information security in 1984. He is a co-inventor of the cipher named MULTI, which is the standard of Japanese Digital Satellite Broadcast System.

In 2001, he moved from Hitachi Ltd. to Tokyo Denki University.



3

My Profile (3)



- (1) Professor, Tokyo Denki University(TDU)
- (2) Director of Cyber Security Institute of TDU
- (3) Former Cyber Security Advisor, NISC (National Center of Incident readiness and Strategies for Cyber Security Information Center, Cabinet Office, Government of Japan)
- (4) Visiting Professor, National Institute of Informatics
- (5) Former General Chair, Japan Society of Security Management
- (6) [Former General Chair of Institute of Digital Forensics](#)

4

Table of contents

1. [What is Digital Forensics ?](#)
2. Early History of Digital Forensics in Japan
3. Activities on Institute of Digital Forensics
4. Digital Forensics Education in Japan
5. Main Cases Involving Digital Forensics in Japan
6. Main Researches in Japan
7. Future Directions



5

What is Digital Forensics ? (1)

- Digital forensics is the process of uncovering and interpreting electronic data. The goal of the process is to preserve any evidence in its most original form while performing a structured investigation by collecting, identifying and validating the digital information for the purpose of reconstructing past events.

The context is most often for usage of data in a court of law, though digital forensics can be used in other instances.

<https://www.techopedia.com/definition/27805/digital-forensics>

6

What is Digital Forensics ? (2)

Forensic Medicine:「法医学」

Technology to clarify information necessary for investigation and trial by using medical knowledge

Murder case



What causes of death? ,
How about a weapon?

Digital Forensics:「デジタル・フォレンジック」

Technology to clarify information necessary for investigation and trial by using information processing technology

Intrusion



What is the intrusion method?
What is the invasion route ?

7

User Candidate of Digital Forensics

- (1) Public Institution
 - (a) Police Agency
 - (b) Prosecution Agency
 - (c) Financial Services Agency etc



- (2) Company

For example, Company uses the DF to keep the evidence of illegal taking out of trade secrets by employees .

8

My meeting with digital forensics

1. When I met DF: around 2002.
2. Trigger: There are many friends of lawyers, and they used the word digital forensics
3. Reasons for starting research (around 2003)
 - (1) In the future, it is expected that most of the data will be digitized
 - (2) It is expected that the awareness of rights will increase and the number of civil lawsuits will increase

=> DF is sure to become important



9

Table of contents

1. What is Digital Forensics ?
2. [Early History of Digital Forensics in Japan](#)
3. Activities on Institute of Digital Forensics
4. Digital Forensics Education in Japan
5. Main Cases Involving Digital Forensics in Japan
6. Main Researches in Japan
7. Future Directions



10

Early History on Digital Forensics in Japan

In 1996: The Japan National Police Agency (NPA) set up a section tasked with the mission of dealing with digital forensic issues triggered by the [Subway Sarin Incident](#).

In 2003 : The first company formed to deal exclusively with digital forensics was established in Japan.

In 2004: The institute of Digital Forensics (IDF) was established.



11

Background

On March 20, 1995, Aum Shimrikyo cult members released sarin gas in Tokyo's subway trains, killing 13 passengers and station workers.

12

Background

In Aum Shirikyo, there were many educated members who have high level knowledge with regards to information technologies.

They used cryptography including public key cipher to protect their data files.

=> Japanese National Police Agency set up the section having the mission to handle the digital forensic issue.

Shokou Asahara
Aum Shinrikyo founder

13

Early History on Digital Forensics in Japan

In 1996, The NPA began efforts to deal with the digital forensic issues related to the Subway Sarin Incident.

In 2003: [The first company formed to deal exclusively with digital forensics was established in Japan.](#)

In 2004: The Institute of Digital Forensics(IDF) was established.



14

Trend on Digital Forensics in Foreign Countries

- ①1984 : Computer Analysis and Response Team was born in US FBI.
- ②1985 : Computer Crime Department started in British Metropolitan Police.
- ③1989 : Michael White developed Forensic Tool named IMDUMP.
- ④ High level commercial Digital Forensic tools such as EnCase and FTK were born in 1990's.
- ⑤1992 : The word, Computer Forensics was used firstly in academic paper by Collier, P.A. and Spaul, B.J.
- ⑥2001 : First DFRWS, which is the conference with concerns to Digital Forensics was held.

http://en.wikipedia.org/wiki/Digital_forensics

15

Table of contents

1. What is Digital Forensics ?
2. Early History of Digital Forensics in Japan
3. Activities on Institute of Digital Forensics (IDF)
4. Digital Forensics Education in Japan
5. Main Cases Involving Digital Forensics in Japan
6. Main Researches in Japan
7. Future Directions

16



Main IDF Activities

In 2004: [The IDF was established](#). The first digital forensic conference, which was called the Digital Forensic Community, was held in December of this year.

In 2006: The Encyclopedia of Digital Forensics was published by Nikka Giren under the supervision of the IDF.

In 2011: The first Digital Forensic Introductory Training hosted by IDF was conducted.

In 2012: The Guideline for Maintaining Evidence (Version 2) was released by the IDF.

In 2015: The Revised Encyclopedia of Digital Forensics was published by Nikka Giren under the supervision of the IDF.



17

Institute of digital forensics(IDF)

The IDF is a non-profit organization (NPO) dedicated to spreading and promoting digital forensics, as well as contributing to the realization of a healthy information technology (IT) society.

IDF membership includes security researchers, digital forensic engineers, people concerned with digital forensic law and law enforcement, as well as digital forensic users.



18

Main Member of IDF at Formation

The screenshot shows the website for 'デジタル・フォレンジック研究会' (The Institute of Digital Forensics). The page lists the following members and their roles:

- 会長 辻井 重男 (情報セキュリティ大学院大学 学長)
- 副会長 安富 亨 (慶應義塾大学大学院法務研究科 法学部教授・弁護士)
- 理事 林 祐一郎 (情報セキュリティ大学院大学 学部長)
- 佐々木 良一 (東京電機大学 工学部 情報メディア学科 教授)
- 高橋 泰夫 (弁護士)
- 赤川 賢洋 (新潟大学法学部 法政コミュニケーション学科 助手)
- 萩原 栄幸 (社)コンピュータソフトウェア著作権協会 技術顧問)
- 海橋 信 (財)東京工科大学 参与)
- 野村 泰貴 (南山大学大学院 法務研究科 教授)
- 佐井 肇哉 (千葉大学 法政学部 助教授)
- 上野 智太郎 (京都大学大学院 工学研究科附属情報センター 助教授)
- 新山 昌純 (国立防衛医療センター 医療情報システム開発研究部 部長)
- 吉川 俊治 (慶應義塾大学大学院法務研究科 医学部 助教授)
- 守本 正忠 (株)UEG 代表取締役社長)
- 石井 正毅 (株)NTTデータ デジタルセキュリティビジネスユニット長)
- 丸谷 修博 (株)フォーカスシステムズ 情報事業推進室 室長)
- 高井 徹 (シー・インサイト・セキュリティ(株) 代表取締役社長)
- 伊藤 一彦 (株)金融システム総合研究所 取締役)
- 佐藤 慶浩 (日本ヒューレット・パッカード(株) 個人情報保護対策室 室長)
- 川向 太郎 (株)情報通信総合研究所 政策研究グループ シニアリサーチャー)
- 丸山 満彦 (監)トーマツ エンタープライズリスクサービス部 シニアマネージャー)
- 黒平 英吾 (財)クモラセキュリティ(株) 専務理事)

General Chair :
 Shigeo Tsujii (Security Researcher)
 (President of Institute of
 Information Security)

Vice Chair :
 Kiyoshi Yasutomi (Lawyer)
 (Prof. of Keio University)

My membership
 number :004
 I was a general chair from 2012 to
 2017.

IDF Membership Growth

	2004	2017	2017/2004
Number of Individual Members	85	259	300%
Number of Corporate Members	25	59	240%



Main IDF Activities

In 2004: The IDF was established. The first Digital Forensic conference, which is called the Digital Forensic Community was held in December of this year.

In 2006: The [Encyclopedia of Digital Forensics was published](#) by Nikka Giren under the supervision of IDF.

In 2011: [The first Digital Forensic Introductory Training hosted by IDF was held.](#)

In 2012: [The Guideline for Maintaining Evidence \(Version 2\) was released by the IDF.](#)

In 2015: The Revised Encyclopedia of Digital Forensics was published by Nikka Giren under the supervision of the IDF.



21

Digital forensics related events in Japan

[Beginning in 2004, Japan-U.S. collaborative investigations on Digital forensic matters started between Tokyo Denki University etc. and Mississippi State University.](#)

In 2005: Digital Forensics was selected as one of the most important 11 security technologies in a report published by the Secretary of Cabinet in Japan.

In 2008: The Fourth Digital Forensic International Conference, which is hosted by the International Federation for Information Processing, Technical Committee 11 (IFIP TC11), was held in Japan.



22

Digital forensics related events in Japan

Beginning in 2004, Japan-U.S. collaborative investigations on Digital forensic matters started between Tokyo Denki University etc. and Mississippi State University.

In 2005: Digital Forensics was selected as one of the most important 11 security technologies in a report published by the Secretary of Cabinet in Japan.

[In 2008: The Fourth Digital Forensic International Conference, which is hosted by the International Federation for Information Processing, Technical Committee 11 \(IFIP TC11\), was held in Japan.](#)



23

Table of contents

1. What is Digital Forensics ?
2. Early History of Digital Forensics in Japan
3. Activities on Institute of Digital Forensics
4. [Digital Forensics Education in Japan](#)
5. Main Cases Involving Digital Forensics in Japan
6. Main Researches in Japan
7. Future Directions



24

Background starting CySec

- The shortage of security experts is a big issue also in Japan.



25

Shortage of Security Field Workers in Japan

Number of Specialists Required (347,000)

Number of Current Security Field Workers (265,000)		Short-fall : (82,000)
Workers (Skilled) (106,000)	Workers (Unskilled) (159,000)	

<http://www.ipa.go.jp/files/000040646.pdf> July, 2014
 IPA: INFORMATION-TECHNOLOGY PROMOTION AGENCY



26

Overview of CySec

- Tokyo Denki University launched a cyber-security education course named CySec in 2015.
- CySec is a course for Security workers and Master course students.
- It is supported by the Ministry of Education, Culture, Sports, Science and Technology (MEXT)



27

Courses in CySec

- 1PF: Cyber Security Infrastructure
- 2CD: Cyber Defense Actual Exercise
- 3IN: Security Intelligence, Psychology, Ethics and Law
- 4DF: [Digital Forensics](#)
- 5MG: Information Security Management and Governance
- 6DD: Secure System Design and Development



28

CySEC

- 1PF: Cyber Security Infrastructure
- 2CD: Cyber Defense Actual Exercise
- 3IN: Security Intelligence, Psychology, Ethics and Law
- 4DF: [Digital Forensics](#)
- 5M: ...nd
- 6D: ...ent

It is a first regular course on digital forensics in a Japanese University.



29

Digital Forensics Curriculum in CySec①

1. Introduction of Digital Forensics
2. Hard disk structure, File system Technologies
3. OS for forensics
4. Forensic work basics
5. Forensic work, Data conservation
6. Forensic work, Data recovery
7. Forensic work, Data analysis①
8. Forensic work, Data analysis②



30

Digital Forensics Curriculum in CySec②

9. Forensic work exercise
10. Network forensic
11. Network forensics exercise
12. DF methods for typical targets①
13. DF methods for typical targets②
14. Law literacy and handling court
15. Future development of digital forensics



In course of 2016, mobile forensics was added instead of DF methods for typical targets②

31

Lecturers

- (1) Prof. Sasaki (Tokyo Denki Univ.)
- (2) Prof. Uehara (Ritsumei Univ.)
- (3) Prof. Yamaki (Tokyo Denki Univ.)
- (4) Mr. Sakuraba (Lawyer)
- (5) Mr. Shirahama (Forensics Expert)
- (6) Mr. Nozaki (Forensics Expert)



32

Education Status

1. In 2015, the CySec course was attended by 54 security field workers and 16 Master course students.
2. Numerous security experts were among the students.
3. Security field workers were sent from police departments, financial services agencies, etc.
4. Based on post-course questionnaire results, students were highly satisfied with our lectures.
5. This course is continued now without support of MEXT.



33

Lecture Scenery of DF



In the digital forensic lecture, practical training of a simulated court based on digital evidence.

The digital forensic lecture was highly appreciated as being organized most.

Picture obtained from Mr. Sakuraba

34

Model Course on Digital Forensics Proposed by Purdue University



Table 4.3 List of Required Courses and Electives

Required Courses	Electives (Specialized Courses)
Introduction to Digital Forensics	Network Forensics
Advanced Digital Forensics	Mobile Device Forensics
Research in Digital Forensics	File System Forensics
Digital Forensics Capstone Course	Anti-Forensics
Thesis or Directed Project	Incident Response
	Digital Law
	Malware Forensics

<http://docs.lib.purdue.edu/dissertations/>

The Development of a Standard Digital Forensics

Master's Curriculum Kathleen Strzempka *Kathleen A. Strzempka,*

kstrzemp@purdue.edu

35

Table of contents

1. What is Digital Forensics ?
2. Early History of Digital Forensics in Japan
3. Activities on Institute of Digital Forensics
4. Digital Forensics Education in Japan
5. [Main Cases Involving Digital Forensics in Japan](#)
6. Main Researches in Japan
7. Future Directions



36

Two Cases Related with Digital Forensics

1. Improper Arrest Case Related to Remote Control Malware
2. Improper Use Case of Shogi Software



37

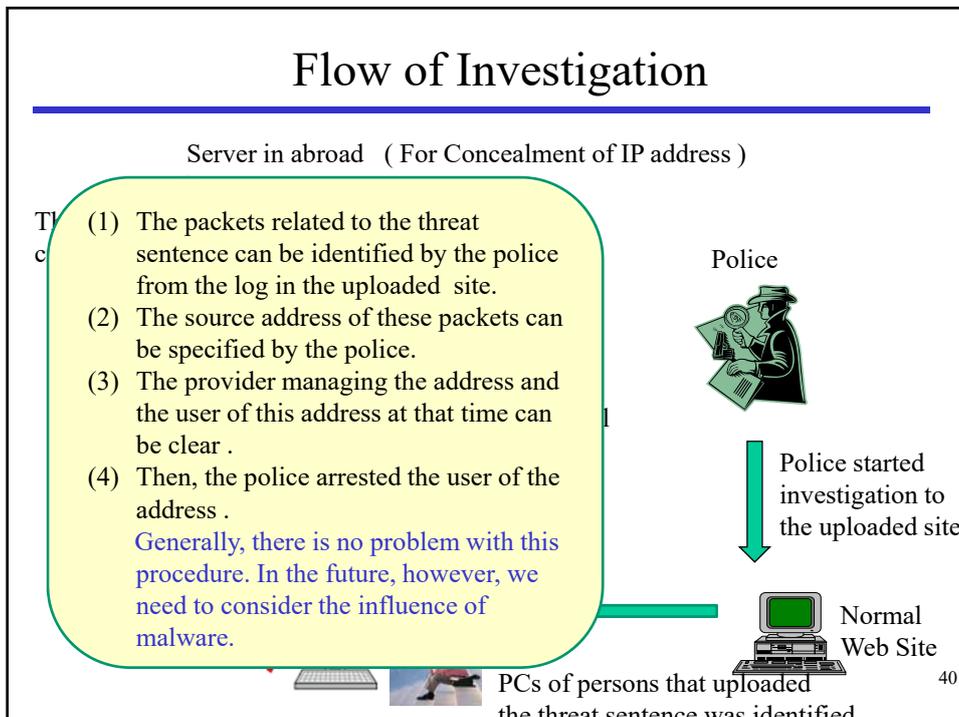
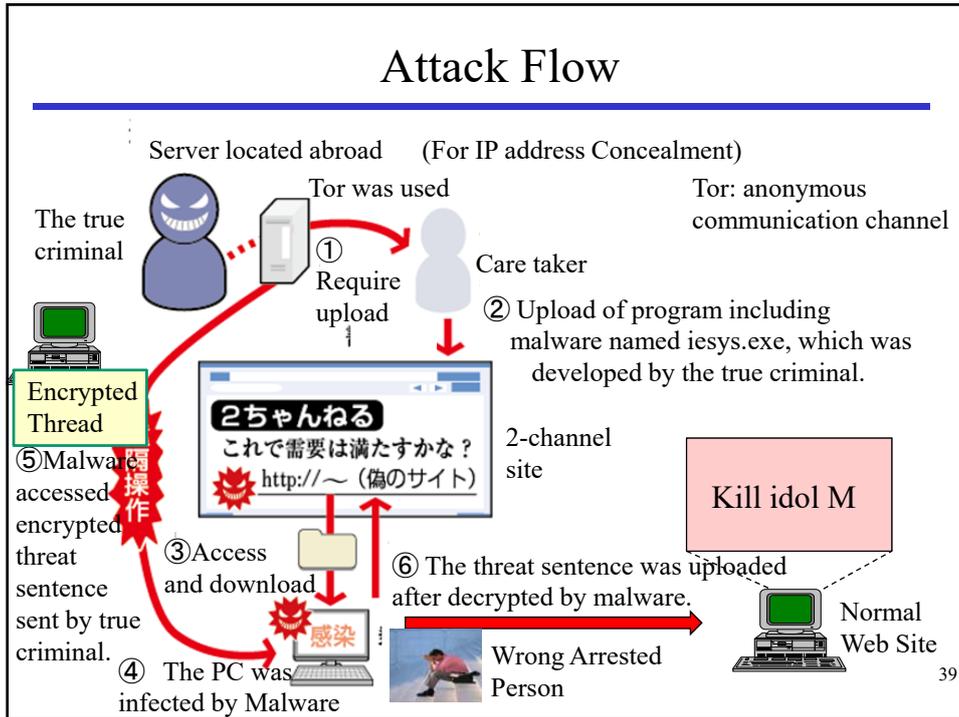
Improper Arrest Case Related to Remote Control Malware

In 2012, four persons were arrested after being suspected of uploading threats such as “Kill idol M” to the website in the Internet.

Later, it became clear that remote control malware in the suspects’ personal computers (PCs) were responsible for the uploading.



38



New Progress

- (1) After then, following message was sent to mass media:

“I have attached a memory chip containing the iesys.exe source program and a text file describing my objectives to a cat on Enoshima Island”

41

Photograph of Enoshima



New Progress

- (2) The cat with a memory chip attached to its neck was discovered by the police.
- (3) At the same time, the police examined image data of surveillance camera in Enoshima showing the memory chip being attached to the cat's neck.

42



New Progress

- (4) A 30-year-old man, hereafter described as “X”, was arrested on Feb. 2, 2013.
- (5) Police announced they had found evidence in the suspect’s PC that showed “X” had accessed Tor around the same time when the malware was uploaded via Tor.
- (6) “X” pleaded not guilty. In his appeal, he stated that he could not write the C# used for iesys.exe.



43

New Progress

- (7) During the trial, the prosecution’s digital forensic expert testified that a piece of the source program remained in the hard disc of the suspect’s PC, thereby providing evidence.

This case marked that the first time deep discussions regarding digital forensics were held in a Japanese court.



44

Results

- (1) After the suspect was released on bail, he held a press conference with his lawyers on May 16, 2014.
- (2) Around the same time, mail from a person who claimed to be the real criminal was sent to mass media. This convinced many people still that “X” was not the actual criminal.



45

Results

- (3) However, a detective who tailed the suspect after his release witnessed him burying a mobile phone on a riverbank.

When the phone was examined, the police discovered same sentence, which the suspect had set to be sent out at the same time as the press conference.

Faced with this evidence, “X” confessed to the crime.



46

Results

- (4) In 2015, the Tokyo District Court condemned 10 crimes including violation of hijacking prevention law and convicted eight years of imprisonment.

Digital forensics has become a very important technology in Japan's courts also.



47

Two Cases Related with Digital Forensics

1. Improper Arrest Case Related to Remote Control Malware
2. [Improper Use Case of Shogi Software](#)



48

Improper Use Case of Shogi Software

1. On October 11, 2016, the Japanese Shogi Federation requests the Miura Kudan to explain at the Executive Committee on suspicion that the Shogi software in smart phones and others was used unfairly.
2. On October 12th, it was decided to suspend the official competition until December 31st.



Shogi is a game similar to chess.
Currently, Shogi software is stronger than professional Shogi player.

49

Improper Use Case of Shogi Software

3. On October 27, it was decided to establish a third party investigation committee.
4. On December 26, the Third-Party Investigation Committee announced that there is no evidence of using Shogi software based on digital forensic survey.



Wikipedia

50

Digital Forensic Survey

1. Digital Forensic Specialized company

Fronteo

2. Electronic equipment which received submission

- One smart phone contacted by Miura Kudan
- One smart phone contacted by wife of Miura Kudan
- Two desktop PCs and one Laptop PC used by wife of Miura Kudan
- One desktop PC and one Laptop PC used by Miura Kudan
- One tablet used by mother of Miura Kudan



https://www.shogi.or.jp/news/investigative_report_1.pdf

51

One of Results Obtained by Forensic Survey of Fronteo

解析対象電子機器	将棋GUIアプリケーション	将棋ソフト	リモートデスクトップアプリケーション	本件4対局中の起動・使用状況
1 三浦棋士が使用していたスマートフォン (以下「本件スマートフォン」という。)	確認されず	確認されず	確認されず	本件4対局中の使用履歴は確認されず
2 三浦棋士の配偶者が使用していたスマートフォン	確認されず	確認されず	確認されず	本件4対局中に多数の使用履歴が存在するが、局対機確認による使用とは判断されない

Fronteo reported that there is no evidence of using Shogi software based on any digital forensic survey.



https://www.shogi.or.jp/news/investigative_report_1.pdf

52

My Comment to the Article in Mainichi Shinbun

① When someone accessed personal computer software, even if he disguise that he is not accessing, the expert can find the evidence which is remained in PC etc.

② PCs and smart phones other than the Miura Kudan were widely used for survey.

↓

Possibility without access to Shogi software is very high.

Jan. 17, 2017

53

Table of contents

1. What is Digital Forensics ?
2. Early History of Digital Forensics in Japan
3. Activities on Institute of Digital Forensics
4. Digital Forensics Education in Japan
5. Main Cases Involving Digital Forensics in Japan
6. [Main Researches in Japan](#)
7. Future Directions

54



Related Main Journals

- Journal of Digital Investigation
- Journal of Digital Forensic Practice
- International Journal of Digital Crime and Forensics
- ADFSL The Journal of Digital Forensics, Security and Law
- International Journal of Digital Evidence
- International Journal of Forensic Computer Science
- Small Scale Digital Device Forensic Journal
- The Journal of Applied Digital Forensics and eDiscovery
- IEEE Transactions on Information Forensics and Security
- Digital Evidence and Electronic Signature Law Review
- International Journal of Cyber-Security and Digital Forensics



55

Articles Related to DF in Japan

- There are no journal specifying digital forensics in Japan.
- We searched [CiNii](#) to find the articles related to “Digital Forensics” in Japan.

[CiNii](#) is a searchable database service containing academic information on articles, books, etc in Japan.



56

Number of Articles According to Year

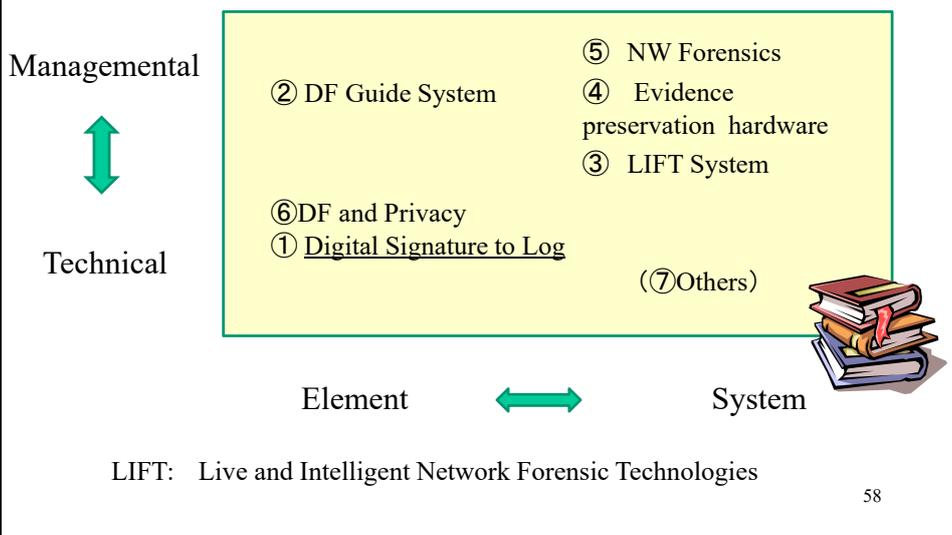
Year	Number of Articles
2006	4
2007	6
2008	11
2009	13
2010	2
2011	7
2012	4
2013	11
2014	7
2015	12
Total	78

Total Number of Articles: 78
Average Number of Articles: ~8

The number is very small, though Japanese papers presented in other countries are not included in these figures.
Our research accounts for the Majority in Japanese digital forensic technology research.

57

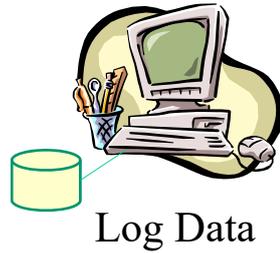
Position of Our Main Researches



58

Background

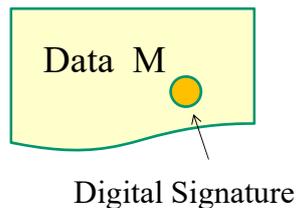
- In recent years, the requirements for preserving important log data as evidence have increased.



59

Basic Scheme and Its Issue

- As a scheme to detect the tampering of digital data, a digital signature scheme is generally used.
- This mechanism is a combination of the public key cipher and the hash function.



$$\text{Sig} = S(h(M))$$

where
Sig: Digital Signature
h: Hash function
S: Public key encryption
using a secret key

60

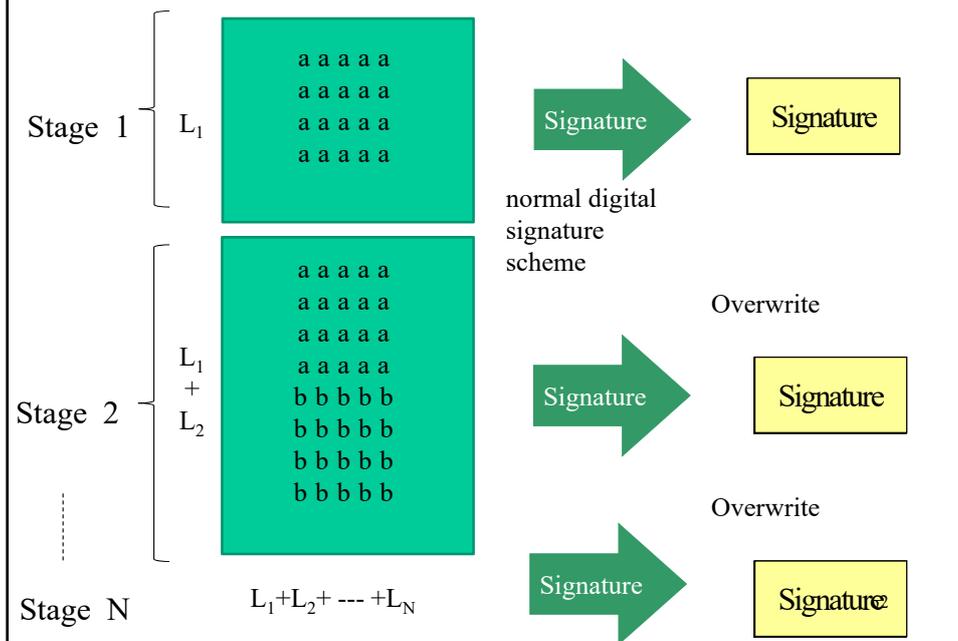
Basic Scheme and Its Issue

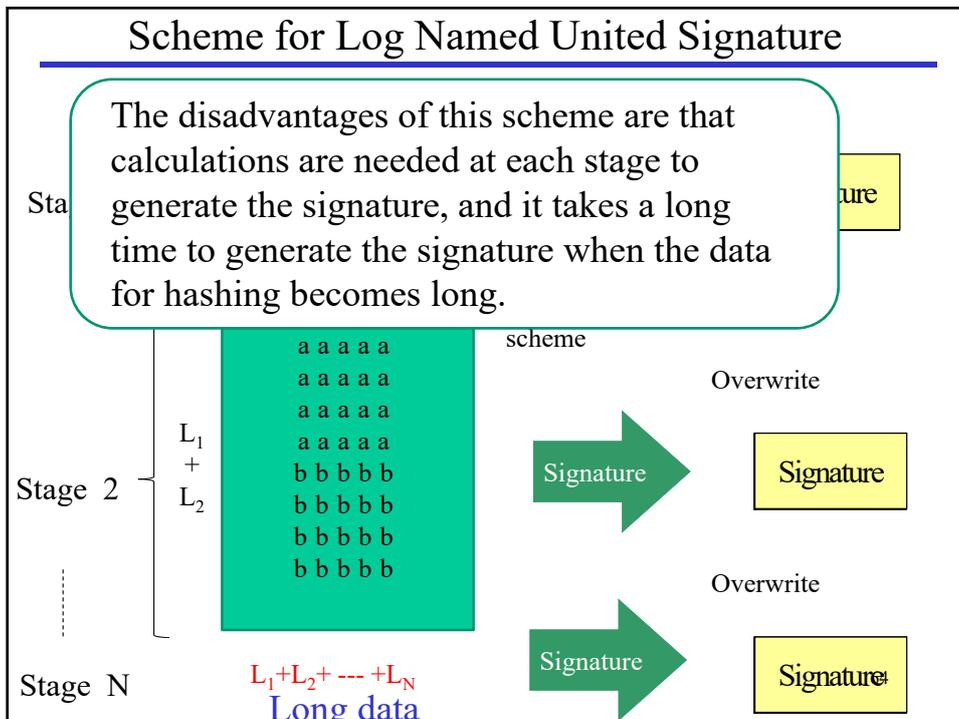
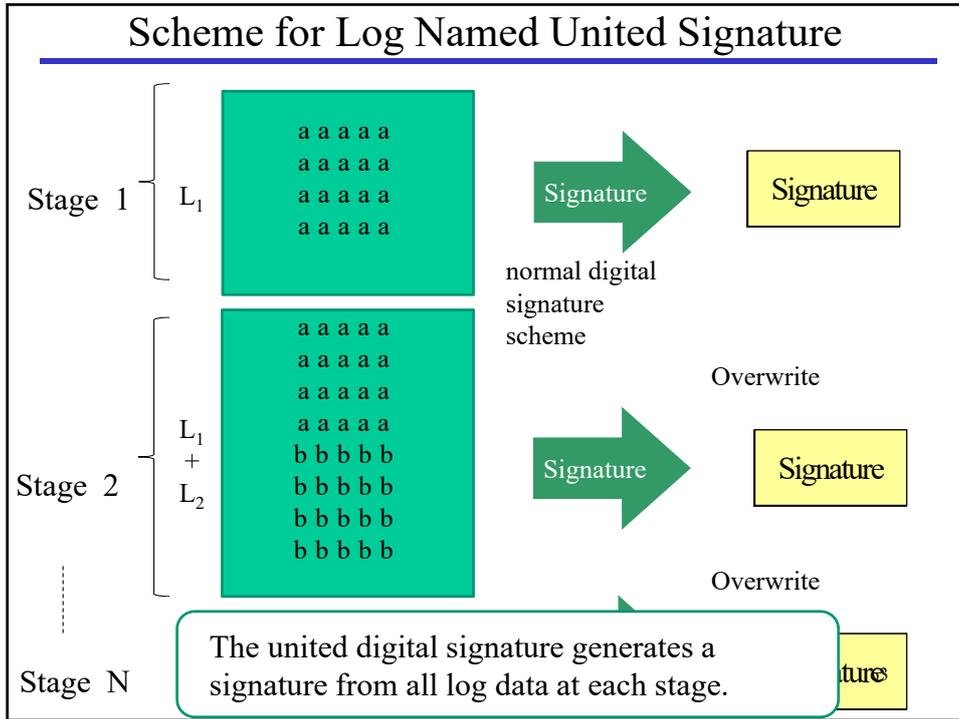
- However, it is impossible to detect log data tampering using a normal digital signature scheme because log data appears intermittently.
- If both the digital data and its related digital signature are deleted together, it is very difficult to detect the deletion in the digital forensics verification phase.

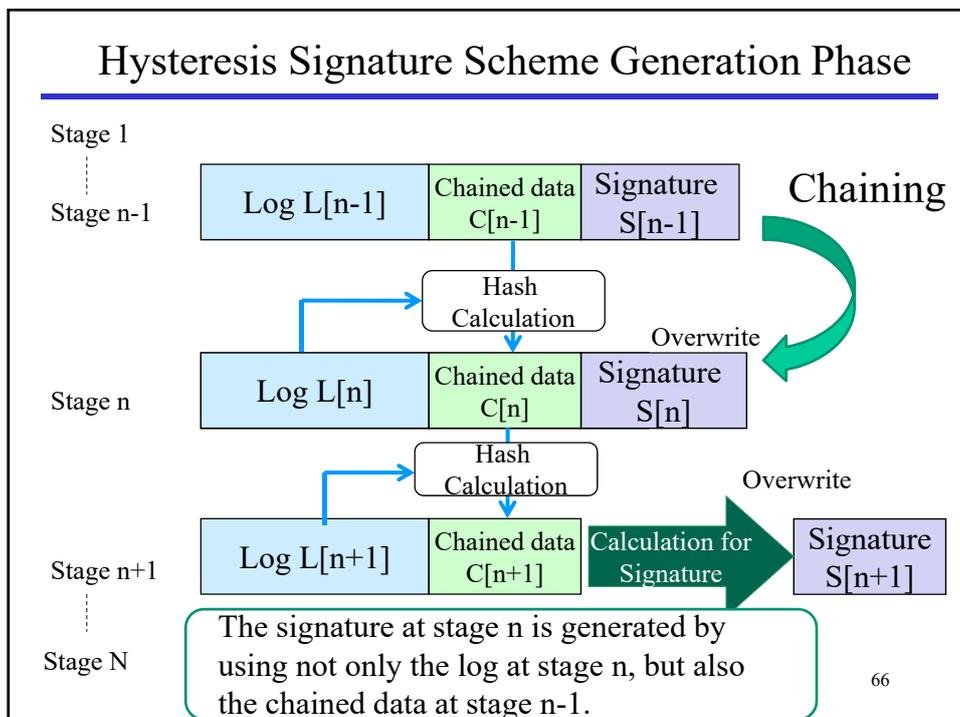
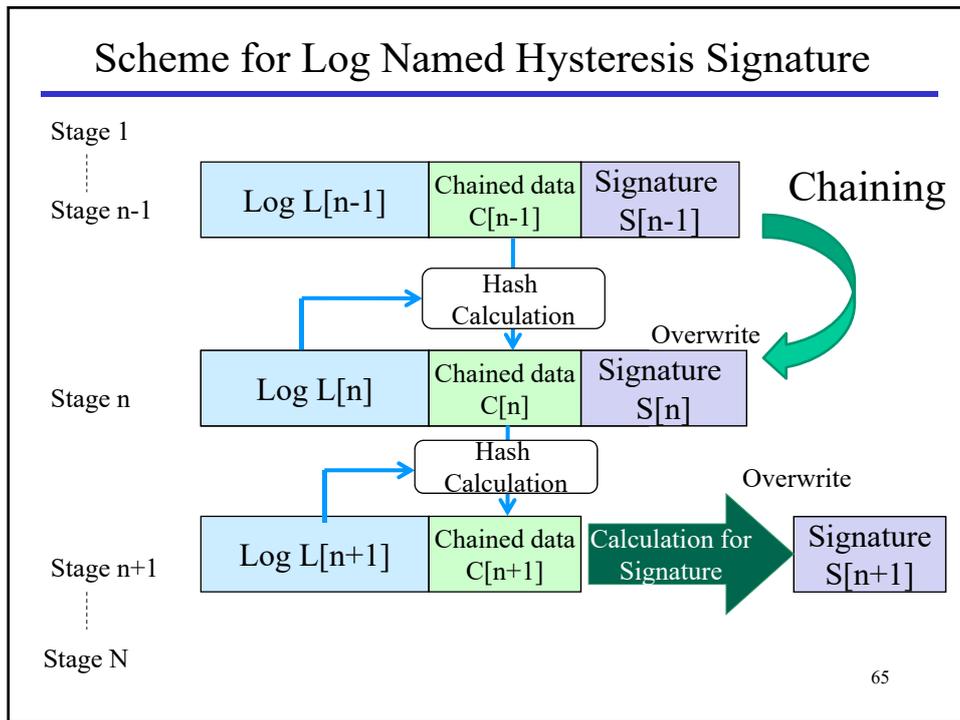


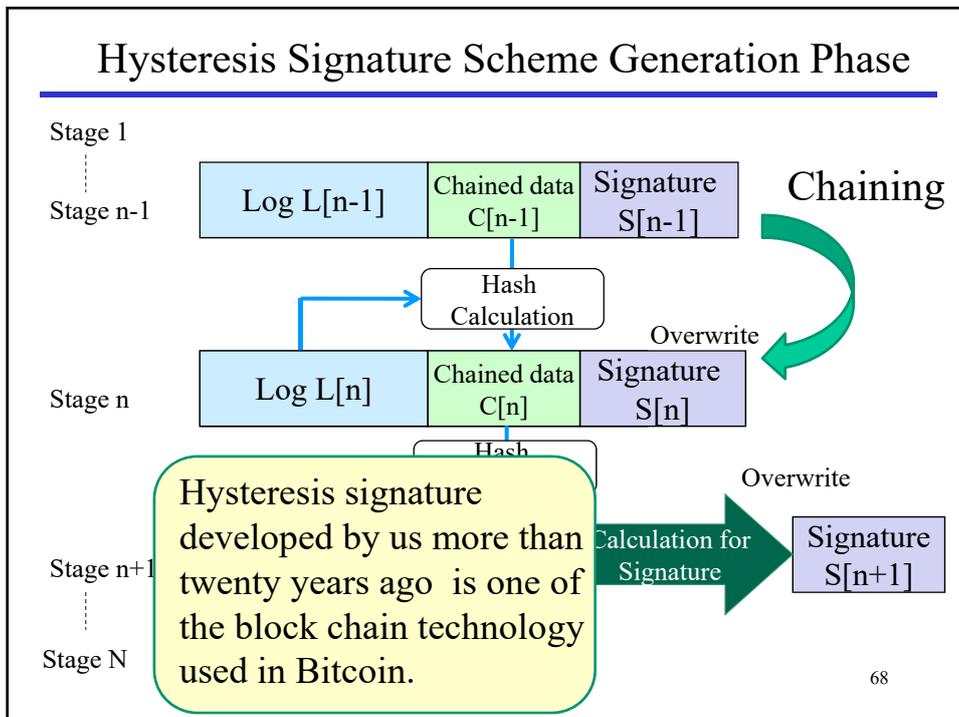
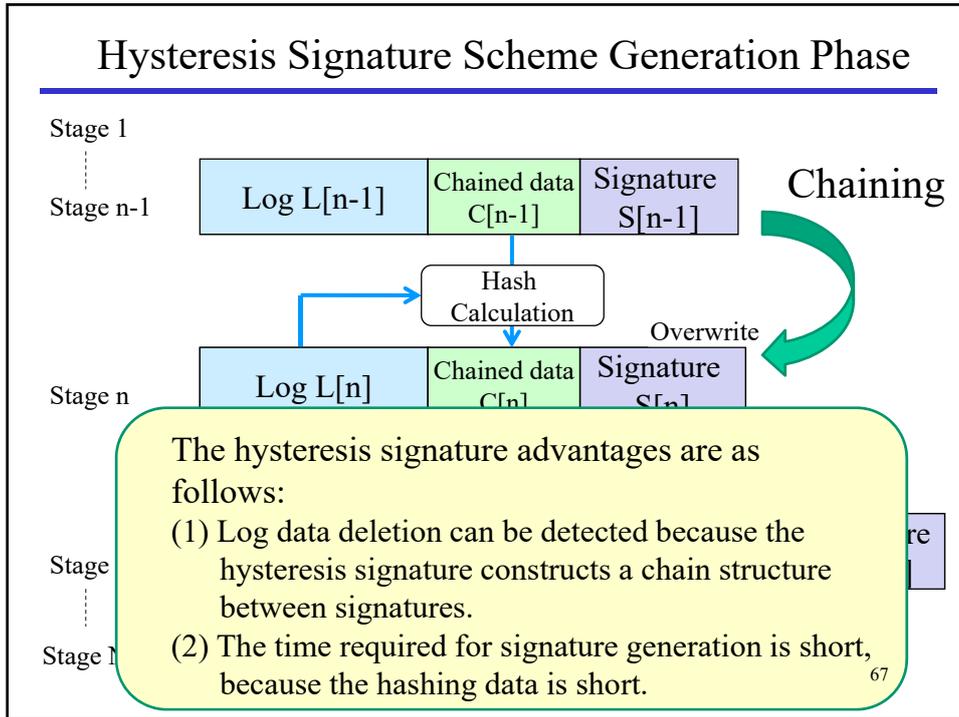
61

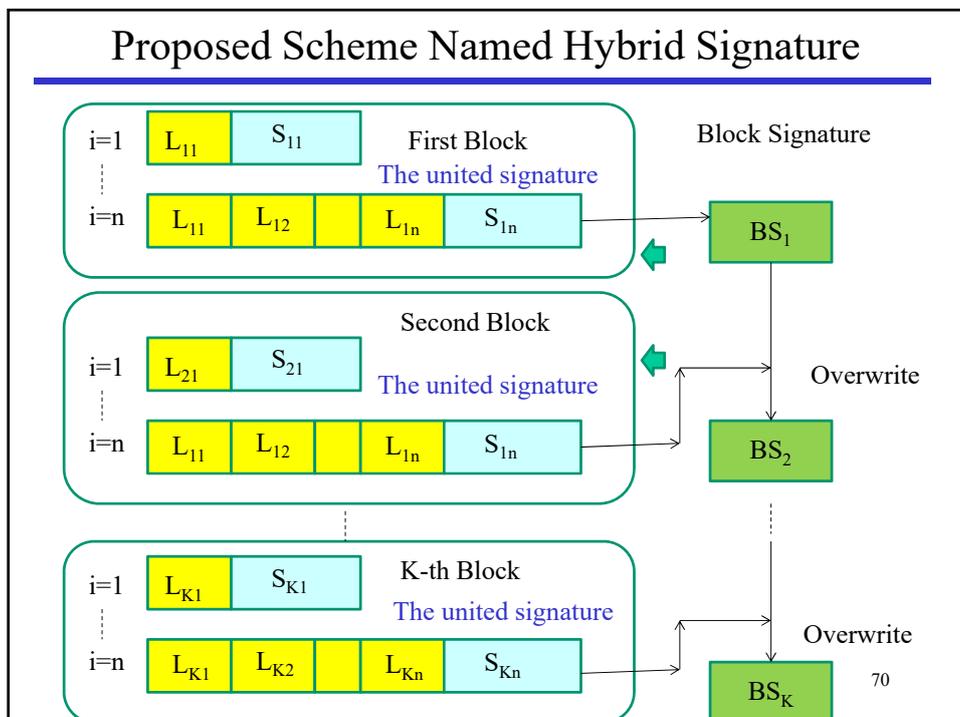
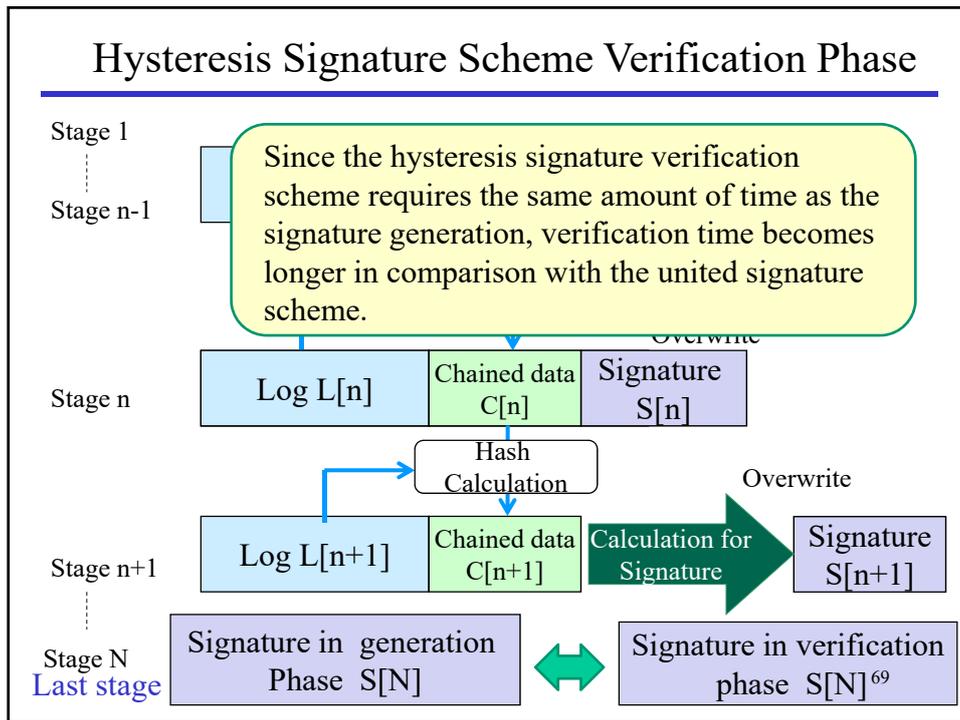
Scheme for Log Named United Signature

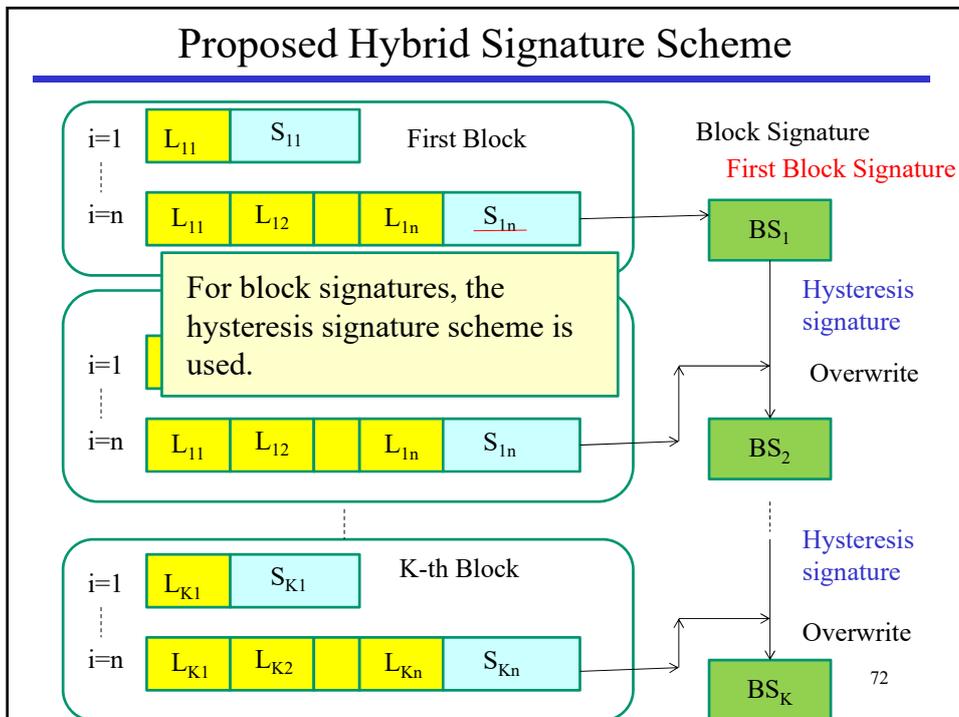
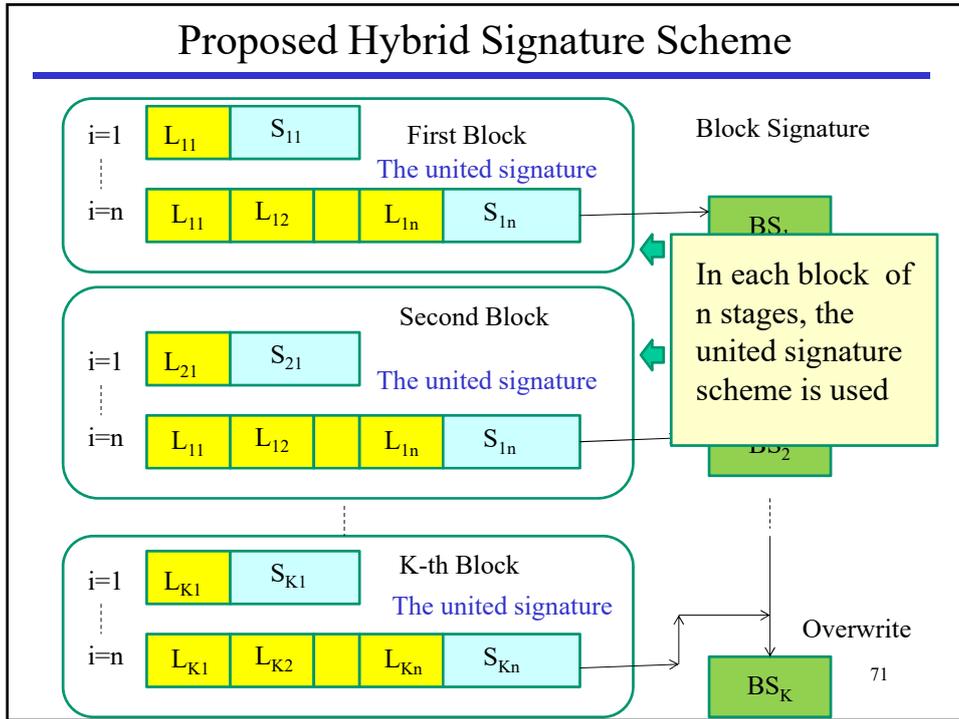




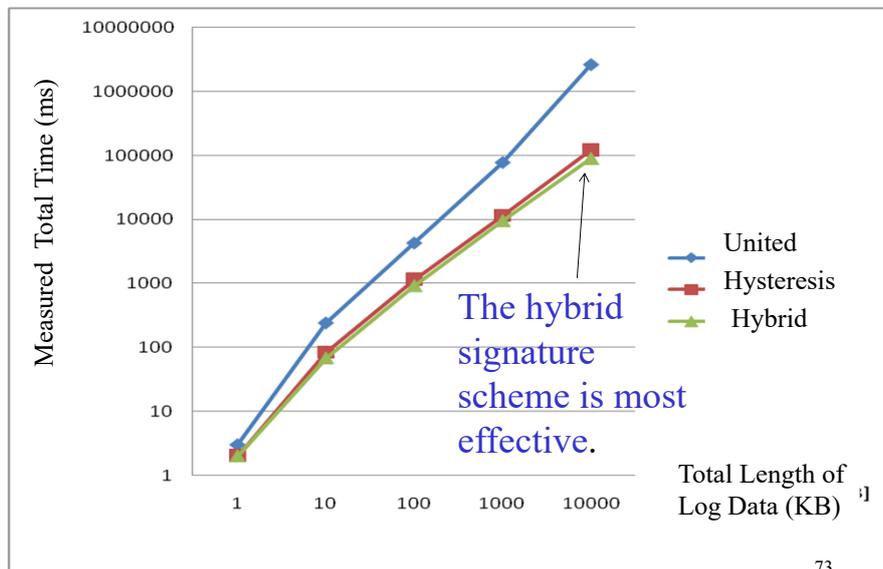








Measured total computation times



73

Related Journal Papers

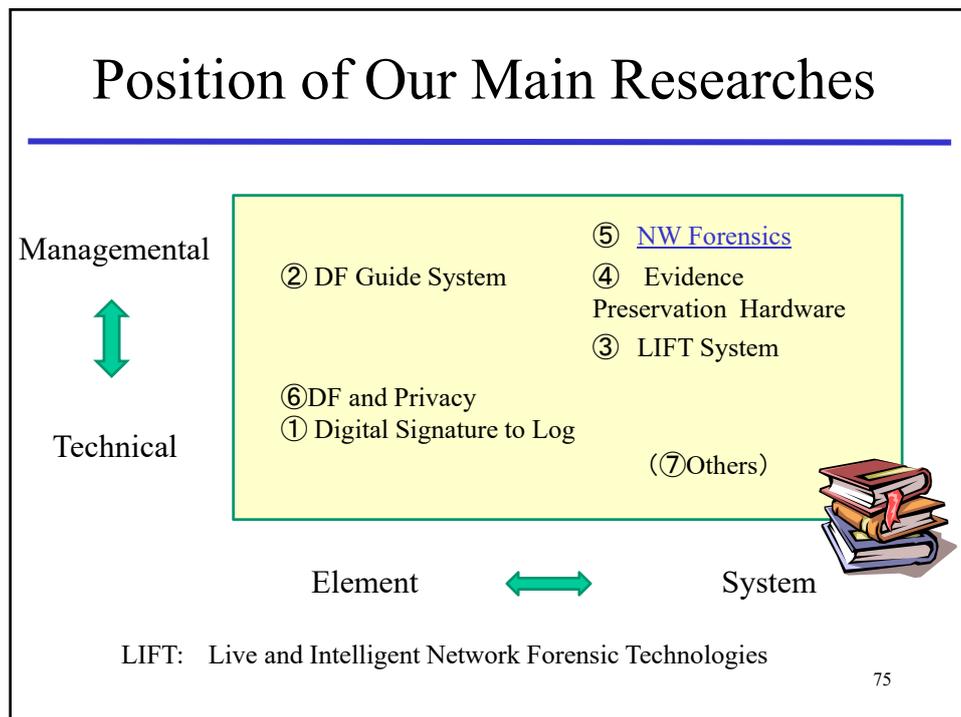


1] Yusuke Ueda, Ryoichi Sasaki et al. "Proposal of Evaluation Method of Hysteresis Signature Scheme Considering Data Loss" Journal of Information Processing Society in Japan. Vol.45, No.8, pp1966-1976, 2004 (In Japanese)

2] Naoki Kobayashi, Ryoichi Sasaki, "Proposal and Evaluation of a Secure and Efficient Log Signature Scheme for Preservation of Evidence", Journal of Japan Society of Security Management Vol. 28, No.2, pp11-21, 2014 (In Japanese)

3] Naoki Kobayashi, Ryoichi Sasaki, "Proposal and evaluation of an evidence preservation method for use in a common number system" international Journal of Electronic Commerce Studies vol.6,no.1,pp51-68, 2015

74



What is network forensics ?

Marcus Ranum defines Network forensics as "the capture, recording, and analysis of network events in order to discover the source of security attacks or other problem incidents".

Marcus J. Ranum
Marcus J. Ranum (born November 5, 1962 in New York City, New York, United States) is a computer and network security researcher.

76

Study Background

- In recent years, cyber attacks via LAN have become increasingly advanced.
- Network forensics become important to identify a cause of unjust communication and to discover the source of security attacks.



77

Study Objective

62.113.232.164	54	49446	>	http	[ACK]	Seq=231	Ack=154	win=131328	Len=0
62.113.232.164	54	49446	>	http	[ACK]	Seq=231	Ack=154	win=131328	Len=0
192.168.137.69	54	http	>	4944					
62.113.232.164	54	49446	>	ht					
62.113.232.164	54	49447	>	F					
192.168.137.255	92	Name	quer						
192.168.137.69	54	http	>	49					
64.4.11.42	363	GET	/	HT					
192.168.137.69	714	HTTP/1.1							
64.4.11.42	54	49437	>	F					
178.250.245.198	66	49450	>	F					
192.168.137.69	66	http	>	49					
178.250.245.198	54	49450	>	F					
178.250.245.198	779	GET	/V7Mc						
192.168.137.69	54	http	>	49					
192.168.137.69	207	HTTP/1.1							

Packet Status

Running processes



While it is possible to identify personal computers engaging in unjust communication by monitoring the packet communication, it is often very difficult to determine the process used by the malware to cause the PC to engage in unjust communication.

78

Study Objective

```
62.113.232.164 54 49446 > http [ACK] Seq=231 Ack=154 win=131328 Len=0
62.113.232.164 54 49446 > http [FIN, ACK] Seq=231 Ack=154 win=131328 Len=0
192.168.137.69 54 http > 49446 [FIN, ACK] Seq=154 Ack=231 win=15680 Len=0
62.113.232.164 54 49446 > http [ACK] Seq=232 Ack=155 win=131328 Len=0
62.113.232.164 54 49447 > http [RST, ACK] Seq=1 Ack=1 win=0 Len=0
192.168.137.255 92 Name query NB WPAD<00>
192.168.137.69 54 http > 49446 [ACK] Seq=155 Ack=232 win=15680 Len=0
64.4.11.42 363 GET / HTTP/1.1
192.168.137.69 714 HTTP/1.1 302 Found (text/html)
64.4.11.42 54 49437 > http [ACK] Seq=1255 Ack=41924 win=65280 Len=0
```

We would like to identify the running process in the PC connected to packet .



Running processes:



79

STUDY OBJECTIVE

To answer the requirement,

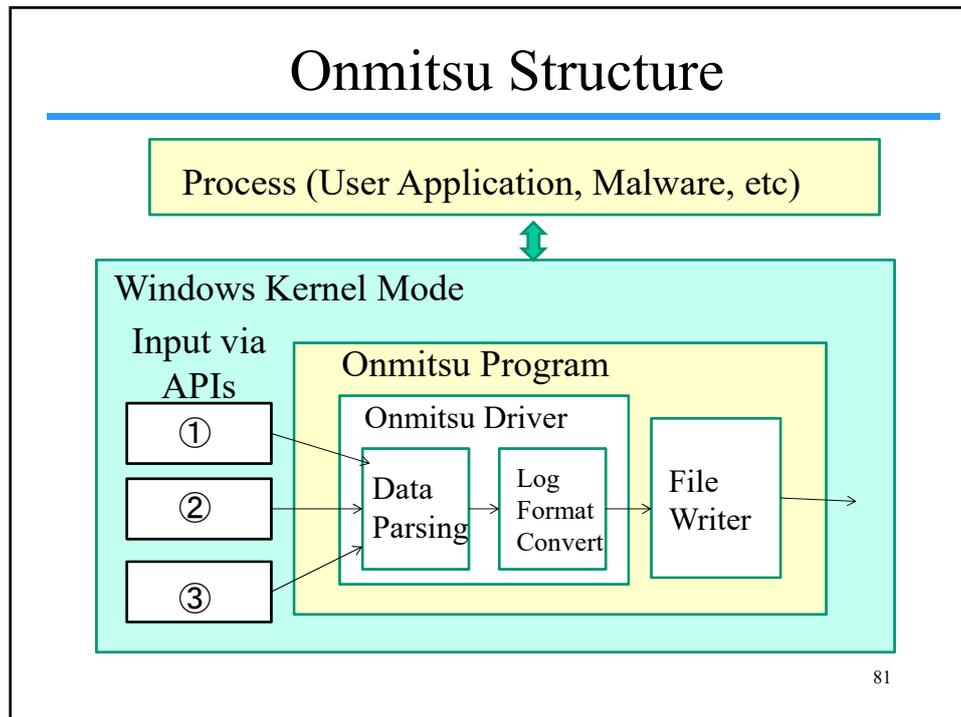
IN 2014, WE DEVELOPED THE LOGGER DRIVER PROGRAM NAMED “ONMITSU”.

“Onmitsu”.is an ancient person who engaged in an intelligence activity.



80

Onmitsu Structure

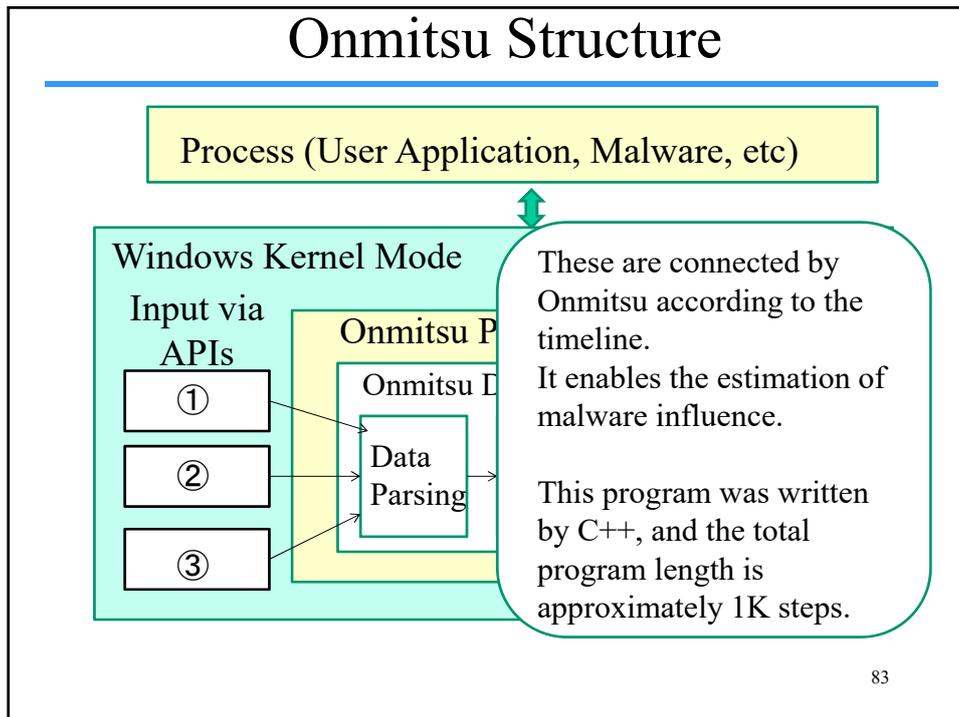


APIs for Input to Onmitsu

- APIs for Input.
 - Windows Filtering Platform(WFP) - ①
It is used to obtain network information
 - PsSetCreateProcessNotifyRoutineEx - ②
 - PsSetLoadImageNotifyRoutine - ③
These are used to obtain process information



Onmitsu Structure



83

Evaluation Items in Experiment

- ① Log accuracy
- ② Log usefulness
- ③ Log volume
- ④ System load



84

Result

- The log from Onmitsu was accurate and useful.
- There were no problems with regards to system load and log volume.
- Onmitsu was transferred to a company and commercialized under the name CapLogger in the company.

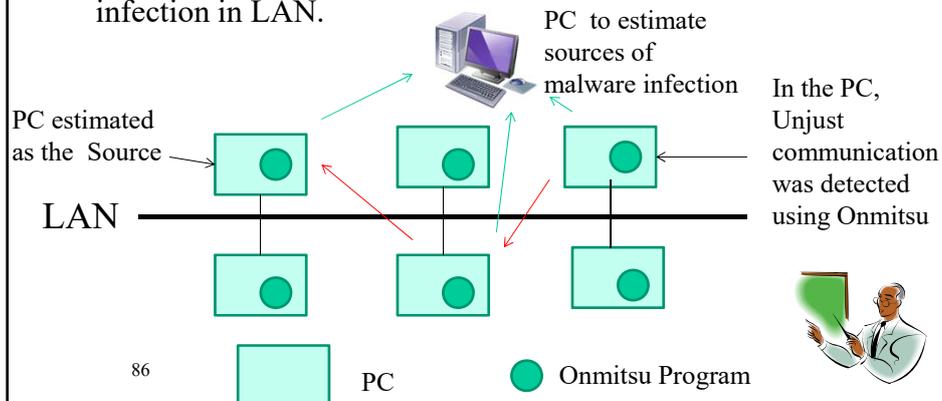


85

*CSV : comma-separated values

Recent Status

- (1) We established a method to estimate sources of malware infection using Onmitsu and ontology in multiple PCs on the LAN.
- (2) Now, we are establishing a method to estimate range of infection in LAN.



86



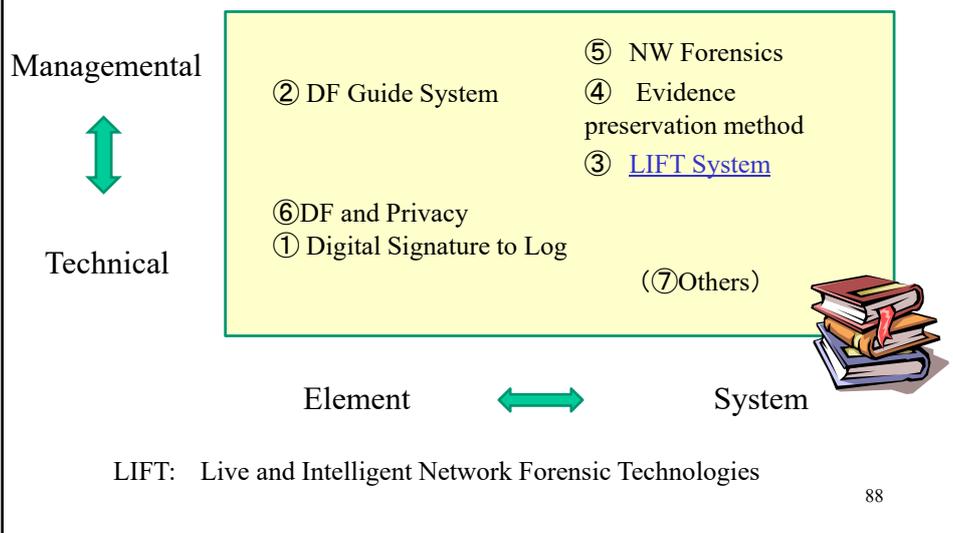
Related Papers



- 1) Satoshi Mimura, Ryoichi Sasaki” Proposal of the Method for Estimating the Cause of Unjust Communication by Using the Network Packets Associated with Process Information” The International Conference on Information Security and Cyber Forensics (2014.10)
- 2) Satoshi Mimura, Ryoichi Sasaki,” Proposal and Evaluation of the Preservation Method of the Network Packets Associated with Process Information” Journal of Information Processing Society in Japan, ,Vol.57,No.9,pp1944-1953,2016 (In Japanese)
- 3) Makoto Sato, Ryoichi Sasaki, Akihiko Sugimoto, Naoki Hayashi, Yoshiaki Isobe “Proposal of a Method for Identifying the Infection Route for Targeted Attacks Based on Malware Behavior in a Network.” Proc. of the Fourth International Conference on Cyber Security, Cyber Warfare, and Digital Forensic (CyberSec2015), IEEE, Oct 2015.
- 4) Makoto Sato, Ryoichi Sasaki,“A Proposal of a Infection Route Detecting Method for Targeted Attacks using Malware Behaviors in the Network” Journal of Information Processing Society in Japan, Vol.58,No.2,pp1-9,2017 (In Japanese)

87

Position of Our Main Researches



88

Background

- Targeted attacks have been increasing year by year



- ▶ It is difficult to perform proper countermeasures against targeted attacks without the assistance of a support system.

89

Background

- SIEM attracts the attention
 - The system combines the functions of security event management and log analysis to provide real-time network forensics.
- However
 - It is difficult to protect attack by using only the SIEM system, because operators need enough knowledge and skill to use the system appropriately.



SIEM: Security Information and Event Management

90

Overview of LIFT Project and System

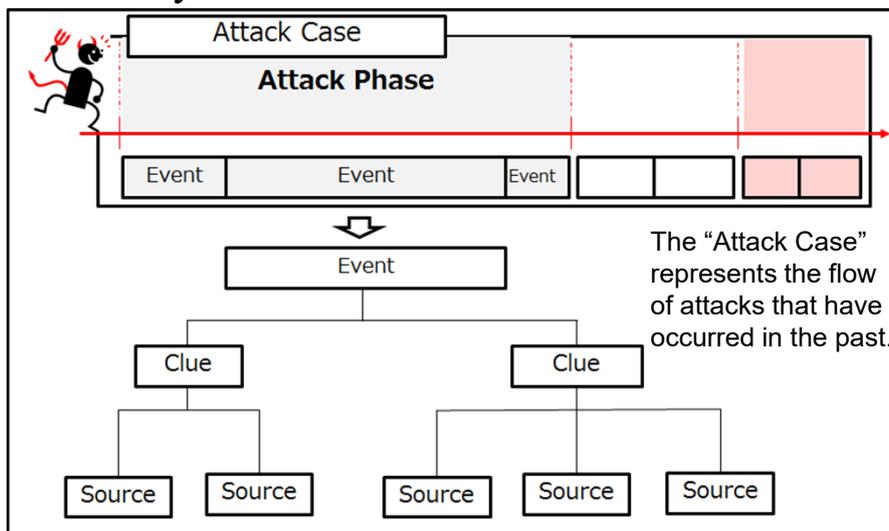
- To cope with the issue, the LIFT project began at the Cyber Security Research Institute of Tokyo Denki University in 2013.
- In the project, we developed the LIFT system having the function of automatic operation using artificial intelligence(AI) and providing proper actions response guidance during incidents



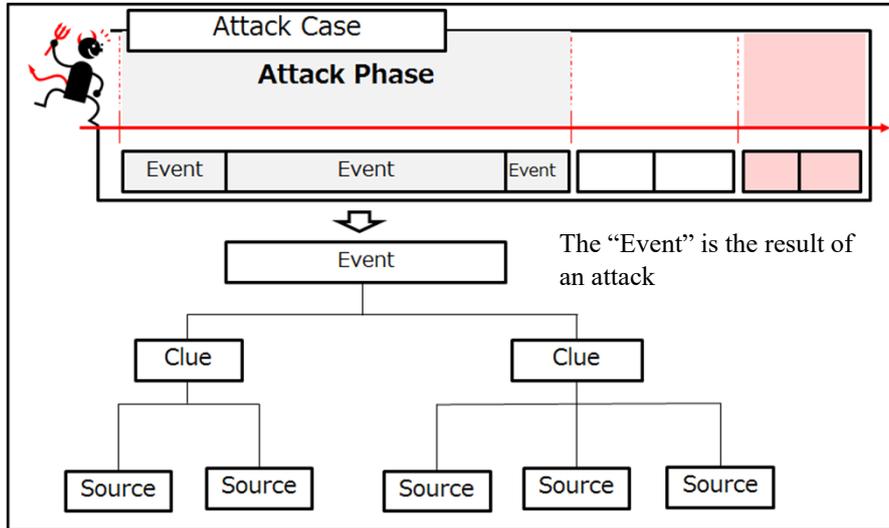
LIFT: Live and Intelligent Network Forensic Technologies

Attack Structure and LIFT System Terms

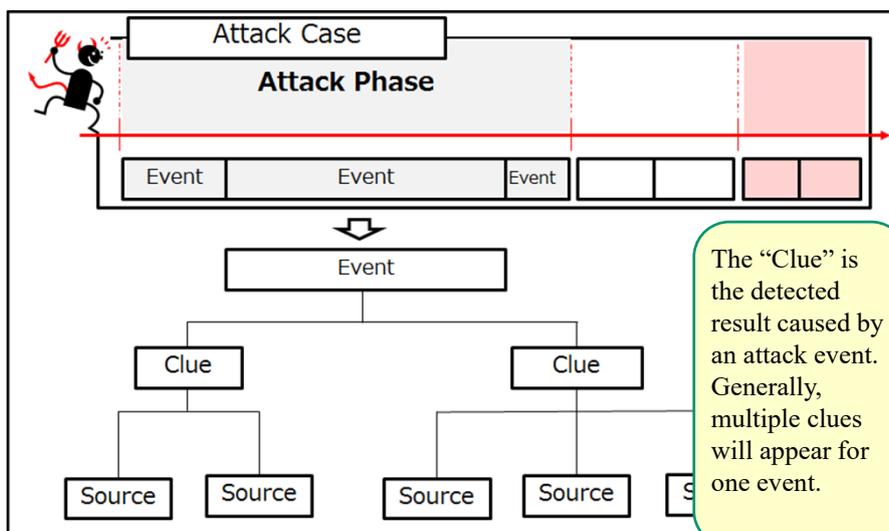
- LIFT system



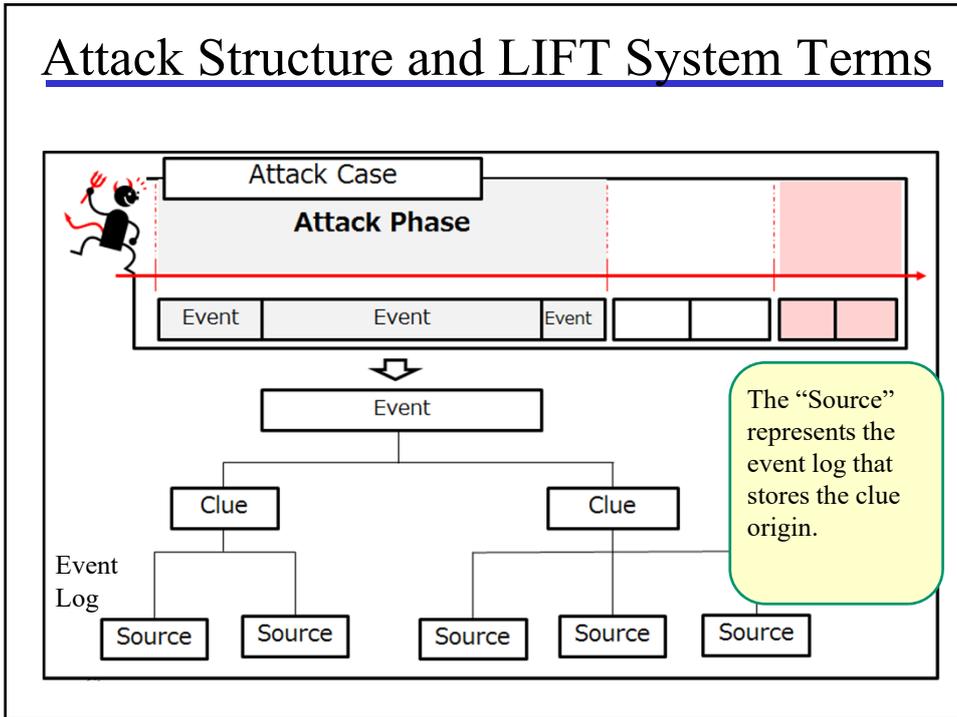
Attack Structure and LIFT System Terms



Attack Structure and LIFT System Terms

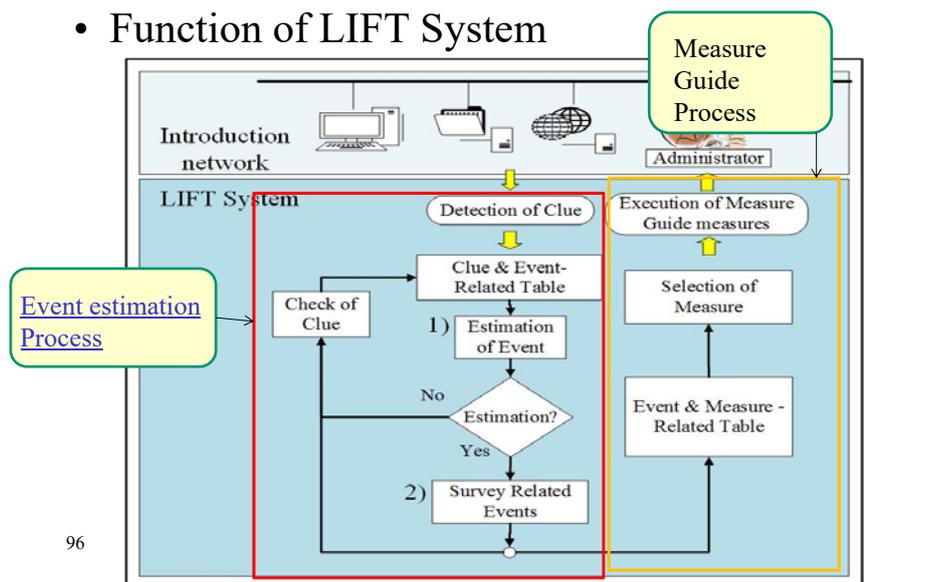


Attack Structure and LIFT System Terms



Overview of LIFT System

- Function of LIFT System



Example of Event and Clue Related Table

	Event	Clue	Proxy Server			
			The execution of the suspicious process	Communication without passing proxy	Using the CONNECT method other than port 443	Long session
Attack Infrastructure Construction	Malware execution	0.3				
	Communication to C&C		0.6	0.6	0.4	
	Download of necessary function for attack	0.4	0.4		0.3	
	Malware collects information of the terminal	0.5			0.2	0.4

Example of Event and Clue Related Table

This table is constructed by experts considering what clues appear, when the event has occurred.

	Event	Clue	Proxy Server			
			The execution of the suspicious process	Communication without passing proxy	Using the CONNECT method other than port 443	Long session
Attack Infrastructure Construction	Malware execution	0.3				
	Communication to C&C		0.6	0.6	0.4	
	Download of necessary function for attack	0.4	0.4		0.3	
	Malware collects information of the terminal	0.5			0.2	0.4

LIFT Project & LIFT System

In operation phase, Clues are observed.
If “communication without passing proxy” is observed, the probability of “Communication to C&C server” is highest.

			Proxy Server			
		The execution of the suspicious process	Communication without passing proxy	Using the CONNECT method other than port 443	Long session	Unnecessary commands to business
Attack Infrastructure Construction	Malware execution	0.3				
	Communication to C&C	0.6	0.6	0.4		
	Download of necessary function for attack	0.4	0.4		0.3	
	Malware collects information of the terminal	0.5			0.2	0.4

LIFT Project & LIFT System

If the value does not exceed the threshold, the other clue related to the event is checked.
In this case “Using the connect method other than port 443” is checked.

			Proxy Server			
		The execution of the suspicious process	Communication without passing proxy	Using the CONNECT method other than port 443	Long session	Unnecessary commands to business
Attack Infrastructure Construction	Malware execution	0.3				
	Communication to C&C	0.6	0.6	0.4		
	Download of necessary function for attack	0.4	0.4		0.3	
	Malware collects information of the terminal	0.5			0.2	0.4

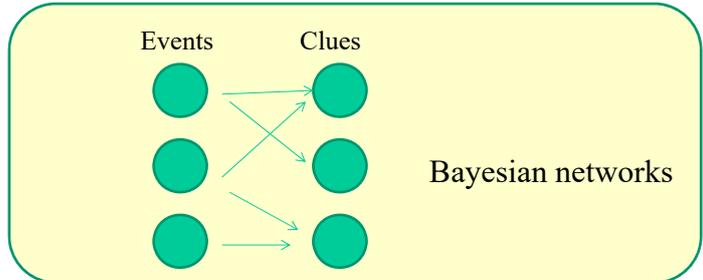
LIFT Project & LIFT System

If the both clues occur, the probability is estimated as
 $P = 1 - (1 - 0.6)(1 - 0.6) = 0.84$
 If the probability exceeds the threshold, the LIFT system guides measure to protect "Communication to C&C".

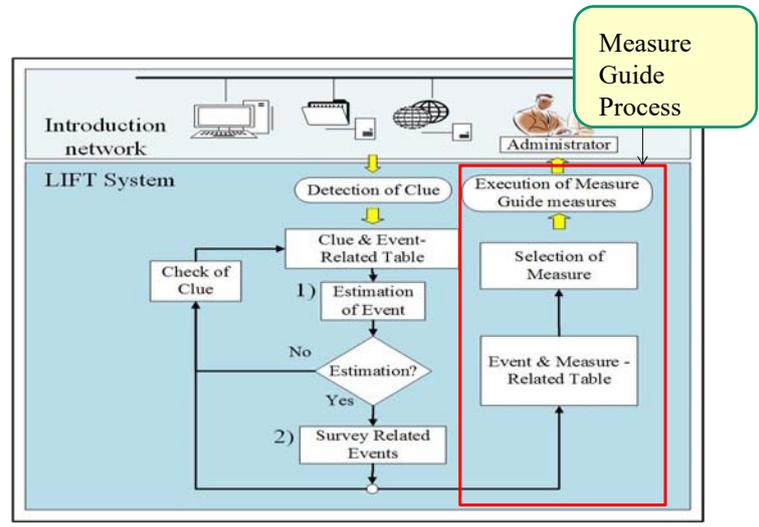
Attack Infrastructure Construction	Event	The ex suspi	Con withou	Using meth	Lo	Unnecessary commands to business
	Malware execution	0.3				
Communication to C&C	0.6	0.6	0.6	0.4		
Download of necessary function for attack	0.4	0.4		0.3		
Malware collects information of the terminal	0.5			0.2	0.4	

Introduction of Bayesian Network

- Now, we use Bayesian networks, which is one of AI technologies, instead of clue and event related tables in order to estimate the event more accurately using the feedback.



LIFT Project & LIFT System



103

Example of Event Measures Related Table

		Measures							
		Blocking the corresponding terminal IP on the router	Blocking of the port on the router	Blocking of inbound communication of the corresponding terminal	Blocking of outbound communication of the corresponding terminal	Isolation of network that applicable terminal belongs	Isolate the corresponding terminal from the network	Stop the appropriate terminal	Process regulation of the corresponding terminal
Attack Infrastructure Construction	Malware execution								
	Communication to C&C	✓	✓		✓				
	Download of necessary factation for attack	✓	✓	✓					
	Malware collects Information of the terminal								

Example of Event Measures Related Table

This table is constructed by experts.

		Measures							
		Blocking the corresponding terminal IP on the router	Blocking of the port on the router	Blocking of inbound communication of the corresponding terminal	Blocking of outbound communication of the corresponding terminal	Isolation of network that applicable terminal belongs	Isolate the corresponding terminal from the network	Stop the appropriate terminal	Process regulation of the corresponding terminal
Attack Infrastructure Construction	Malware execution								
	Communication to C&C	✓	✓		✓				
	Download of necessary factation for attack	✓	✓	✓					
	Malware collects Information of the terminal								

Example of Event Measures Related Table

If “communication to C&C” is identified as event, these three measures are recommended by LIFT system.

		Measures							
		Blocking the corresponding terminal IP on the router	Blocking of the port on the router	Blocking of inbound communication of the corresponding terminal	Blocking of outbound communication of the corresponding terminal	Isolation of network that applicable terminal belongs	Isolate the corresponding terminal from the network	Stop the appropriate terminal	Process regulation of the corresponding terminal
Attack Infrastructure Construction	Malware execution	↑	↑		↑				
	Communication to C&C	✓	✓		✓				
	Download of necessary factation for attack	✓	✓	✓					
	Malware collects Information of the terminal								

Development of Proto Program

- We developed the LIFT proto program written in JAVA in about 2K steps.



107

Example of Proto Program GUI



108

GUI in the case that a clue was observed.

Application experiment

Purpose:

- Confirm the usefulness of the LIFT system
- Determine whether the LIFT proto program meets the LIFT system requirements.



109

Results of Application Experiments



Experiment

- We prepared six attack events
- Each pseudo attack was launched in the experimental environment 10 times
- The experimental results were compared against estimated attack results

Results

Six attack events were able to be detected without fail using this proto program.

110

Recent Status

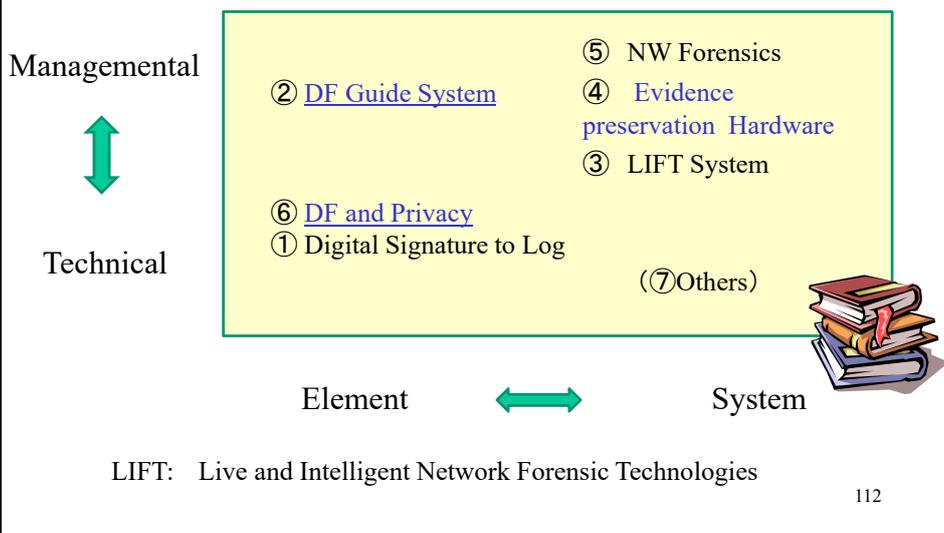


1. Although we were able to identify events that occurred in the past, it was difficult to identify new type events. To cope with this issue, a multi agent approach is being introduced.
2. The function to estimate the sources of malware infection using Onmitsu and ontology in multiple PCs on the LAN is introduced to LIFT system

Recent research status will be presented in CSS2017 organized by Information Processing Society in Japan.

111

Position of Our Main Researches



Overview and Related Paper

② DF Guide System

It is a system that guides correspondence with a smartphone etc. to a first responder for an incident.



Takamichi Amano, Tetsutaro Uehara, Ryoichi Sasaki,
 “Proposal and Development of Guideline Total Support
 System for Digital Forensics” Journal of Information
 Processing Society in Japan, Vol.56, No.9, pp1889-
 1899, 2015 (In Japanese)

113

Overview and Related Papers

④ Evidence Preservation Hardware

Study to make PC as tamper resistant as smart card by introducing simple improvement of hardware and hysteresis signature scheme



- Yuki Ashino, Ryoichi Sasaki, “Development and evaluation of a digital forensic system using Security Device and Hysteresis Signature”, ” Journal of Information Processing Society in Japan Vol.49, No.2, pp. 999-1009, 2008 (In Japanese)
- Keisuke Fujita, Yuki Ashino, Tetsutaro Uehara, Ryoichi Sasaki, “” Proposal of DF system with boot control function against unauthorized programs ” Journal of Information Processing Society in Japan, VOL.51, No.9, pp1507-1519, 2010 (In Japanese)
- Takashi Shitamichi, Ryoichi Sasaki, “Technology of Federated Identity and Secure Loggings in Cloud Computing Environment”, International Journal of Electronic Commerce Studies Vol.5, No.1, pp. 39-62, 2014 doi: 10.7903/ijecs.1157,2014

114

Overview and Related Papers

⑥ [DF and Privacy](#)

Technology to make it possible to ensure both evidence and privacy



(1) Mitsuyuki Takatsuka, Ryoichi Sasaki “Proposal of the e-Discovery System for Sanitizing Disclosure Information and for Securing Evidence” , Journal of Information Processing Society in Japan, Vol. 49, No.9 pp3191-3198, 2008 (In Japanese)

(2) Shuhui Hou, Siuming Yiu, Uehara, Sasaki et al,「A Privacy-Preserving Approach for Collecting Evidence in Forensic Investigation」International Journal of Cyber Security and Digital Forensics (IJCSDF) (Vol.2,No.1pp70-78) 2013

115

Table of contents

1. What is Digital Forensics ?
2. Early History of Digital Forensics in Japan
3. Activities on Institute of Digital Forensics
4. Digital Forensics Education in Japan
5. Main Case Involving Digital Forensics in Japan
6. Main Researches in Japan
7. [Future Directions](#)



116

Future Direction

1. The importance of digital forensics will increase year by year also in Japan.
2. We would like to increase the number of digital forensics experts, including researchers.
3. Personally, I would like to focus primarily on
 - (1) Network Forensics
 - (2) Live Forensics



117

Thank you for your attention



118

Session 3:
Data Models
(Chair: Tomoki Yoshihisa)

Non-Model Approach based Process Improvement Method

†Akihiro HAYASHI and Nobuhiro KATAOKA

†Quality Assurance Group, Onosokki Co.,Ltd, Japan
Interprise Laboratory, Japan
pixysbrain@gmail.com

Abstract - In order to improve quality, cost, and delivery time (QCD) of system development, process improvement has been widely introduced in Japan. However, even when they attain Process Model goal like the capability maturity model integration (CMMI) Level 3, no enough actual effect has been confirmed yet. Therefore, effective process improvement approach is expected. In the paper, we propose non-model approach based process improvement method that identify the process area to improve that align to the organization's goal. We apply this method to realize improvements in a real case process, and the results confirm the actual effectiveness with respect to the QCD.

Keywords: Non-Model Approach, Process Improvement, Defect Reports Analysis

1 Introduction

In order to improve the quality and productivity of the system development, process improvement has been introduced. Process improvement is defined as the realization of a reduction in the need for rework in system development, and an attempt to increase the QCD of system development projects. Both the model approach and non-model approach are well-known methods in process improvement methodology.

The model approach is a method that was introduced by CMMI[1], or ISO/IEC 12207[2] and/or 15504[3], and aimed to achieve the conformance that is expected of process management. The basis for such a process model is the idea that in any organization worldwide, it is common for the minimum required processes to be performed. The model approach includes modeled approaches such as the CMMI's IDEAL approach and ISO/IEC 15504's Geese approach.

On the other hand, the non-model approach does not adopt a process model, but instead focusing on improving individual issues pertaining to each organization. Not only are multiple processes discussed at the same time, such as attaining the maturity level, but unique processes are identified and improved to realize the business goal and impeding the achievement of the organization. In the non-model approach, the total quality management (TQM), goal question metric (GQM), and plan, do, check, act (PDCA) approaches are known.

Recently, there has been much focus on realizing process improvements using the model approach, but it is more realistic to adopt a non-model approach when attempting to solve specific issues pertaining to individual organizations. In addition, in the case where no specific improvement effects, such as QCD improvements, are confirmed even though CMMI stage level 3 has been attained, the concept of the non-model

approach should be introduced recursively, and there is a need to tackle process improvement using the non-model approach.

In this study, we attempt to establish a process improvement method using a non-model approach on the premise of system development using a waterfall model. The procedure first visualizes the magnitude of rework in each process by using the defect reports developed during the project life cycle. Next, we quantitatively identify the work products of the upstream process that have result in the largest rework. Next, we improve the process by adding the necessary practices to the process of development standard in the organization. When we apply the proposed method to an actual project and examine the effect for 3 years, the rework ratio, which was originally 51%, had improved to 42%. The organizational objective test defect ratio improved from 42% to 37%. We have got a prospect that this method is effective.

In a prior study in the field of software process improvement, Fukuyama et al.[4] proposed a tool named SPIS (A Software Process Improvement Support System) that supports process improvement. However, this study focuses on the use of tools in the category of assessment. Sakamoto et al.[5] analyzed the development process formally, and they quantified the effort of the amount of reduction achieved quantitatively, then showed the concrete benefit. Thereby, the method motivated the stakeholders of process improvement. Tanaka et al.[6] has reported cases that motivated the stakeholders of process improvement by describing the current processes of the organization and estimating the improvement effect. The authors of [7] proposed a method that identifies the process that needs to be improved by following the upstream process in accordance with the ISO/IEC 15504.

However, no prior studies reported that focus on the non-model approach, which identifies and improves processes area that cause reworks by analyzing upstream defect reports.

The remainder of this paper is organized as follows. In Section 2, we explain the concept and rework of the current process improvement method, after which we describe the issues to be solved. In Section 3, to solve the issues, we propose a method that visualizes the rework process by analyzing and improving the upstream defect reports. In Section 4, we explain a case study where this proposed method is applied to real-world organizations that have already achieved CMMI maturity level 3. And we then conclude the paper in Section 5.

2 Hypotheses and Issues of the Reworked Process

In this section, we present an overview of the current process-improvement activities. Then, we describe the hypothesis and issues to be solved using the non-model process-improvement approach.

2.1 Process Improvement Approaches

Recently, model approaches have been used to realizing process improvements. As a proof of the effectiveness of the model approach, graphs are presented that show how the QCD has improved in accordance with level up maturity level 1-2-3[8].

However, the model approach does not guarantee improvements in the QCD where CMMI and/or other approaches have been introduced. The graph presented as evidence is just one incident. More importantly, there are cases that report no specific improvement in the QCD effects for organizations in which CMMI was introduced and achieved maturity level 3[9].

The process improvement does not directly improve the QCD, but reduces the "rework" that was originally unnecessary. Then, this resulted in the improvement of the QCD.

2.2 What is a rework?

In order to explain what is a rework, let us quote the project cost classification in system development proposed by Bill Curtis in Figure 1[10].

In system development, the project cost can be classified as an implementation cost and quality cost. The implementation cost is the total cost to manage a project, and includes the preparation of a project plan, progress management, reporting, and the engineering cost for manufacturing. Next, the quality cost is the cost to improve the quality of the product to be provided. Furthermore, quality costs are classified as compliance costs and noncompliance costs. The compliance cost is the sum of the costs related to evaluations, such as reviews, tests, and audits for quality-control purposes, and preventive cost such as training, procedures, tools, and data classification.

Meanwhile, noncompliance is the cost of debugging, correction, retesting, etc., which occur because of the cost of activities to ensure compliance. Because noncompliance costs correct defects that are noted in conmliance activities, there is

no noncompliance cost if there is no noncompliance. According to Bill Curtis' research, as a result of investigating leading global companies such as IBM/TRW/NASA-SEL/HP/Raytheon, 30-40% of a project's cost is the noncompliance cost.

Even if the productivity itself is not expected to improve, the removal of this noncompliance cost by process improvement is expected to significantly improve the QCD. In this research, we define the work generated due to noncompliance in Figure 1 as "rework."

2.3 Hypothesis of Process Improvement to Reduce Rework

A rework can be classified into two types: Noncompliance due to defects before the upstream process and noncompliance due to defects in the process. For example, when a defect is detected by a source-code review, if it was wrongly coded owing to the use of an incorrect design document, it becomes noncompliance due to a defect before the upstream process. When the correct design document is used, if it is an error in the code, it is a noncompliance that is due to a defect in the process. In order to reduce the reworks, it would be prudent to improve the upstream process that resulted in the noncompliance defect in downstream processes.

In this research, we focus on defect reports in order to reduce reworks due to defects in the upper stream process. The defect report describes defects that are detected during the process. The contents of the defect reports are mainly based on basic information such as the date, the person in charge, the defect registration number, and the target system. In addition, the defect detection process, the defect outline, the generation location, the treatment outline, and the defect incorporation process are included.

Here, if the defect-detection process and the mixing process coincide with each other, it is a noncompliance due to a defect in the current process. On the other hand, if the detection process and the mixing process do not match, it can be regarded as a defect caused by the upper stream process. The basis of process improvement is root-cause management, and improving the upstream process can result in greater benefit by improvements in the reduction of rework.

From this perspective, we hypothesized that it is possible to reduce the rework by tracing back to the upstream process that caused the defects using the defect reports; we can identify and improve the processes that are more beneficial to process improvement in the downstream process.

2.4 Issues to be Solved

In order to establish a process-improvement method that is based on a non-model approach, the following two issues should be solved.

1. Establish procedures to identify processes that caused rework.

The process-generating rework is different for each organization. We have not established the method to identify and improve the process of the upstream process that caused the rework in individual organizations.

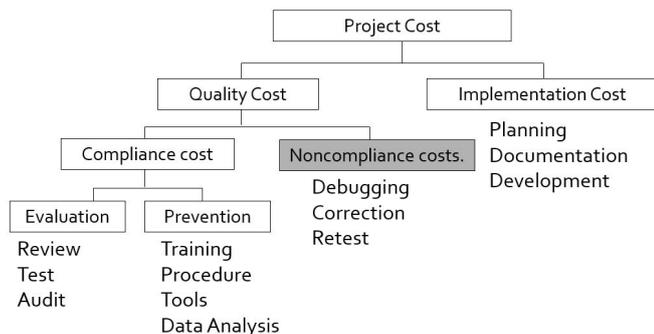


Figure 1: Classification of project costs

2. Establish procedures that contribute to the realization of organizational objectives.

An organization's expectation for process improvement should be to meet some organizational goals. In many cases, there may be an individual goal other than improving QCD which accompanies the reduction of rework. For this reason also, we chose the non-model approach. Procedures for process improvement to contribute to achieving such individual organizational objectives are not yet established

3 Implementation Method of Process Improvement Activities Using Non-model Approach

In Section 3, to solve the issue described in Section 2.4, we propose the process improvement implementation procedure using the rework data of defect reports. The proposed method consists of five phases, the setting of organizational goals for process improvement, the visualization of the rework effort, the identification of work products in the upstream process to be improved, the root-cause analysis for embedding non-compliance, as well as additional practices to eliminate root causes. This situation is shown in Figure 2.

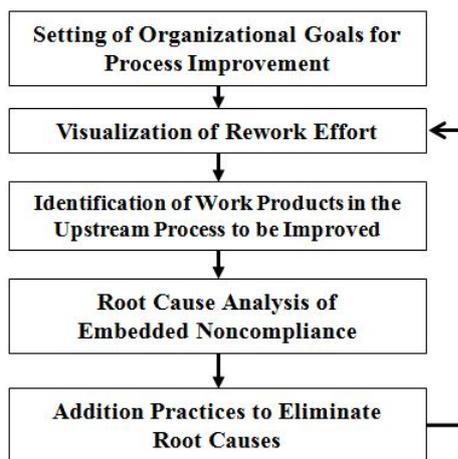


Figure 2: Implementation method for process improvement in non-model approach

3.1 Setting of Organizational Goals for Process Improvement

First, we set up organizational goals for process improvement. At the same time, an index to measure the organizational goal quantitatively should be determined.

Examples of organizational objectives include achieving a given maturity level, improving the quality and process performance objectives (QPPO), improving productivity, reducing market outflow defects, and so on.

The indexes used to measure the organizational objectives quantitatively are achievement requirements that assess whether organizational objectives have been achieved. If in the case where an organizational goal is to reduce the market outflow

defects, the quantitative index should be "the ratio of the number of market outflow defects will be 10% lower than the previous year."

3.2 Visualization of Rework Effort

We calculate the effort ratio for each process in the whole project, and then we visualize the effort expended for each process. In this case, the effort is defined as the modified effort described in the defect reports.

3.3 Identification of Work Products in the Upstream Process to be Improved

In the waterfall-development life cycle, work products that are created in a certain process become the input to the next process. Defects that occur in the downstream process are caused by defects in work products from the upstream process. The reason for the occurrence of some problems in the design process may be because of some element of trouble in the requirement-definition document.

We first pinpoint the upper stream process the defect contamination based on the amount of rework effort, which is described in Section 3.2. The work products made in each process are documented as a work product list at the time that the configuration-management plan is developed. These documents should include work products that caused defects in the downstream process. Considering the unique circumstances of each organization, work products that contain defects are identified logically.

3.4 Root Cause Analysis of Embedded Noncompliance

The reasons for which work products are incomplete are related to cases where the work products themselves were defective, where it is difficult to highlight the defects by peer review, and where the defects are not detected only by peer review. This may be the case where an intention in the upstream process is not transmitted correctly or when similar defects that had to be corrected at the same time are missed. This would be difficult to determine in the review. In such a case, we will follow the work products and processes that caused the upstream process.

3.5 Addition Practices to Eliminate Root Causes

In organizations in which process improvements are introduced, their activities are carried out according to the "organization's set of standard processes (OSSP)". The OSSP is classified in the process area, and is described as the process that produces a specific kind of work product. Therefore, we can identify a process that is performed according to the result of the work product.

By identifying work products that resulted in a defect in the downstream process, we can identify the process that is performed by examining the OSSP. Therefore, we can improve the process by adding practices to eliminate the cause of the defect embedded in the work product.

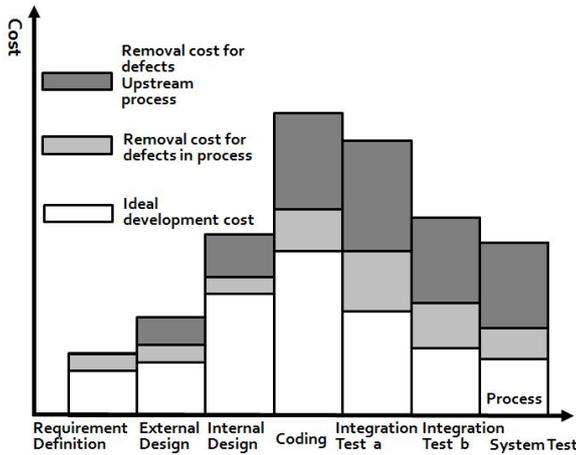


Figure 3: Development Lifecycle and Effort Ratio

The above is the proposed method employed in this research. However, even if one practice is added to a process that has already been established and operated, it is difficult to have a notable effect. More importantly, to contribute to the achievement of organizational objectives, multiple process improvement methods should be simultaneously utilized.

In this research, if defects due to the influence from upstream process account for more than 10% of the defects, the cycle of 3.2 to 3.5 for this proposed method shall be continued.

4 Application Evaluation

In Section 4, we can evaluate the effectiveness of the proposed method by case study the method proposed in Section 3.

4.1 Case

Company A has been working on CMMI activities for 10 years, and achieved CMMI maturity level 3 four years ago. Company A adopts the waterfall model’s development lifecycle, which consists of a requirements definition, external design, internal design, manufacturing, integration test A, integration test B, and system test. Integration test distinguishes integration test A, which performs to satisfy external specifications and internal specifications and Integration test B targets external specifications at interfaces between subsystems.

4.2 1st Round

The organization goal was set as "Left Shift by Reduction of Test Defect Ratio". Left shift is defined as an activity to prevent from rework in the downstream process by detecting larger defects in the upstream process. The test defect ratio defines the number of detected defects in the test process vs. the entire life cycle upon completion of the manufacturing process.

When Company A introduced this proposed method, the test defect ratio was 42%. The index used to measure the

organizational objective quantitatively is "the Test Defect Ratio was Declined for Three Consecutive Years from the Start Point". Figure 3 shows each process on the vertical axis and the effort on the horizontal axis.

Company A carries out a peer review in the upstream process against the requirement definition document, external design document, internal design document, and source code. In the downstream process, the tests are conducted in integration test A and integration test B, and the system test. Then defect reports are respectively documented.

We draw noncompliance costs for defects in the upstream process as dark, incompatible costs for defects in the current process as light, and the ideal development cost as white in Figure 3. The sum of the noncompliance costs for defects before the upstream process and the nonconformance cost for the current process was 51%.

When summing the defect reports, 23% of the dark color area was embedded in the external design process. In the external design process, a system architecture design document, a screen design document for software, and external and internal interface design documents are made. Among them, the architectural design documents are reused, and the screen design document is already agreed with the customer, so we suspect that the defect is embedded in the interface design document.

Next, in order to analyze the root cause of the incompatibility embedded in the interface design document, we conducted investigations from three points of view: 1) Input to the interface design document, 2) completeness of the interface design document itself, and 3) the absence of a peer review of the interface design document. The incompatibility between the interface document and other classified work products is significant.

There is a sub project manager (PM) in charge of electricity, machinery, and software. When we receive a change request from customer, it is necessary to simultaneously modify the interface design document among the sub-PMs. However, some interface design documents will be missed in the sub-PMs. This causes defects in the downstream process. We concluded the root cause of the trouble was the change-management process.

Company A also introduced a change-management process, but often changed the source code suddenly in response to change requests, and sometimes did not revise the interface design document consistently.

A configuration control board (CCB) is not an essential practice in CMMI. Company A arbitrarily performs CCB. Therefore, Company A incorporated practices to perform CCBs based on the judgment of the PM and sub PM when a customer’s requirements changed.

In addition, as there is a possibility that the impact analysis itself is inadequate even after performing the CCB, we also added practices to number the traceability matrix from the requirement definition document to the interface design document. If the traceability matrix is assigned a number, it can also be useful for analyzing the influence range, and it is possible to prevent omissions to corrections of similar parts in the downstream process.

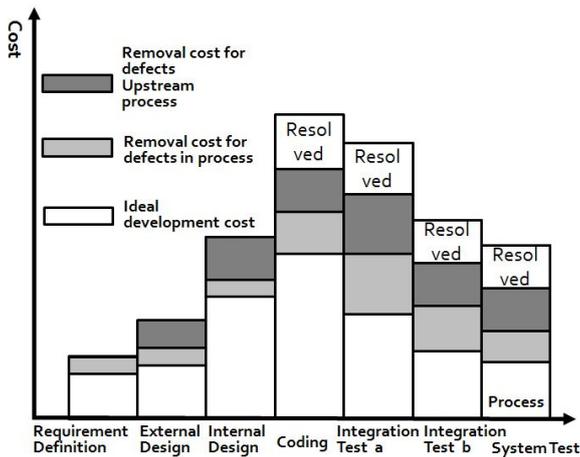


Figure 4: Development lifecycle and effort ratio (2nd)

4.3 2nd Round

Figure 4 is a graph that was created for the visualization of rework after the completion of the 1st round of process improvement proposed in this paper. Because the height of the bar graph signifies the effort ratio of each process, it is the same as in the 1st round. "Resolved" is the area that is expected to be resolved in the 1st round of the proposed method.

After we excluded the portion referred to as "resolved," 12% of the dark area was in the internal design process. This means that even in the second round, more than 10% of the defects were concentrated within a process. Therefore, we proceeded to the next step.

Software design documents and state-transition diagrams are primarily created in the internal design process. In the embedded development, the software design document is patterned beforehand, and defects other than careless mistakes are not often embedded. Therefore, it was inferred that there is a defect in the state-transition diagram. Based on an actual examination, defects were included in the interface between the newly developed part and the diverted part in the derivative development. When a state-transition diagram contains a defect, there should be a notification of a mistake when the state-transition diagram is developed. However, because it is difficult to complete a complicated state-transition diagram, we determined that the main reason of the defect was the absence of peer-review.

In Company A, the implementation rate of the peer review was decided according to the scale of work products. In this case, the peer-review implementation rate for documents such as design documents was 100%, and for transition diagrams and source codes, it was around 30%, centering on "dangerous places" based on the PM judgment.

Therefore, in the peer review of the state-transition diagram, when the reviewer reached 30% the exit criteria, they should terminate the peer review. For this reason, the part including the defect may have deviated from the peer-review target range.

In other words, 30% of the peer-review implementation-rate was not included in the range that the PM judged as "a potentially dangerous place" was the root cause. Therefore, a

new practice was added such that the PM and sub-PM must negotiate the validity of the peer-review scope and prepare the minutes of the meeting.

There are three peer review methods: the buddy method, team method, and reading-out method. The buddy method is performed by two persons, a creator and reviewer. The team method is a method in which multiple reviewers conduct an individual preliminary review. The reading-out method is a walk-through method in conference format. It does not include a preliminary review, and the producer of the work reads the review subject at the review meeting, and explains the work product in order from the beginning to the end. The peer review becomes strict in this order, and instead of requiring many attempts, the defect detection rate increases.

Company A adopts the buddy method when anything is specified, but in the case of "newly developed parts" that are in the "dangerous place", we added the practice to the review with the reading-out method.

4.4 3rd Round

Similarly, in the 3rd round, we created a graph corresponding to Figure 4, and we visualized the classification of incompatibility using color coding. However, because the defects before the pre-process did not exceed 10%, we decided to end the cycle of this proposal.

4.5 Application Result

In this section, we evaluate the effectiveness of the proposed method by using the result of applications over the three-year period.

1. Have procedures been established to identify and improve the process that caused the rework?

Although the proposed method focused on reworks that are due to upstream defects, for comparison with the case where CMMI was introduced, it is necessary to include reworks due to defects of upstream processes and due to in-process defects, as in Figure 5. The rework effort ratio, which was originally 51%, improved to 42%. An improvement effect of approximately 9% was confirmed three years after introducing this proposed method.

2. Has an improvement procedure that contributed to realizing organizational objectives been established?

Figure 6 shows the test defect ratio over three years after introducing this proposed method. The test defect ratio improved from 42% to 37%. It can be said that the confirmed improvement is a significant effect.

5 Conclusion

In this paper, we proposed a process improvement method based on a non-model approach that appears to contribute to the achievement of each organization's business goals.

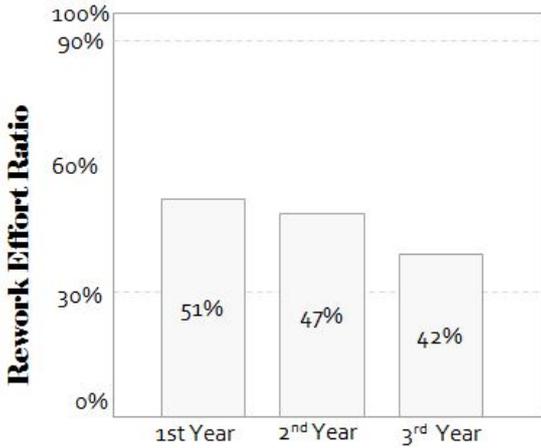


Figure 5: Rework effort ratio

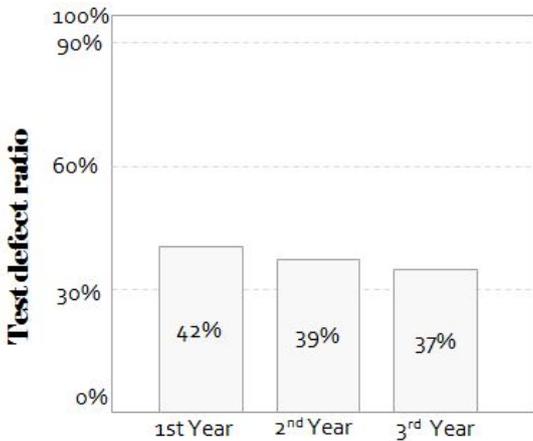


Figure 6: Test defect Ratio

The company A that was referred to in this paper did not confirm the concrete effect while attaining the CMMI maturity level, but in the three-year period after adopting this proposed method, we confirmed the effect of realizing the goal and reduction in the number of reworks. We therefore believe that this proposed method is effective.

To further improve the proposed method, it is necessary to change some subjective judgments in the procedure to make them more objective. The defect reports are sometimes created during the project activity, the embedded upstream process defects are "surmised". Even when multiple factors causing defects are considered, it is sometimes difficult to express them accurately in defect reports.

REFERENCES

[1] SEI, CMMI for DEV Ver1.3, SEI, CMU, 2011
 [2] ISO/IEC 12207, Information Technology Software life cycle processes, JIS X 0160:2012
 [3] Software Process Seminar, Receiving the standard announcement of TR, Japan Standards Association, Information Technology Standardization Center, 2000
 [4] Shun-ichi Fukuyama, Shuu-ichi Miyamura, Hideo Tak-

agi, Ryoji Tanaka, A Software Process Improvement Support System SPIS, IEICE Trans. Inf. Syst.,Vol.E83-D No.4 pp. 747-756, (2000)
 [5] Keishi Sakamoto, Toshifumi Tanaka, Shinji Kusumoto, Ken-ichi Matsumoto, and Tohru Kikuno, An Improvement of Software Process Based on Benefit Estimation, In the Trans of the IEICE. PP. 740-748, (2000)
 [6] Tanaka T. Sakamoto K., Kusumoto S., Matsumoto K., and Kikuno T.: "Improvement of Software Process by Process Description and Benet Estimation" 17th Int'l Conf.e On Software Engineering, pp. 123-132, (1995)
 [7] Akihiro Hayashi and Nobuhiro Kataoka, Identification Method in Software Process Improvement Areas, IEEJ Transactions on Electronics, Information and Systems 128(7), 1231-1241, 2008-07-01
 [8] James Herbsleb, David Zubrow, Dennis Goldenson, Will Hayes, Mark Paulk, "Software quality and the Capability Maturity Model," Communications of the ACM, Volume 40 Issue 6, June 1997, pp 30-40
 [9] J. Herbsleb, A. Carleton, J. Rozum, J. Siegel, and D.Zubrow, Benets Of CMM-Based Software Process Improvement: Initial Results, Software Engineering Institute, CMU/SEI-94-TR- 13, 1994
 [10] Akihiro Hayashi, "Process Improvement Method by Non-Model Approach ", IEICE, SWIM2016-13, vol. 116, no. 341, pp. 11-18, 2016 Dec

Derivation of a Map of Variables in a Loop Structure

Kozo Okano[†], Shinji Kusumoto[‡], and Yukihiro Sasaki[‡]

[†]Faculty of Engineering, Shinshu University, Japan

[‡]Graduate School of Information Science and Technology, Osaka University, Japan
okano@cs.shibshu-u.ac.jp, kusumoto@ist.osaka-u.ac.jp

Abstract - Program analysis enables us to understand the behavior of the program. Analysis of loop structure is, however, difficult in general. In order to solve such a problem, an existing technique derives a map between variables using regression analysis on a data obtained by multiple executions of the program (run history). When we analyze a complicated loop using the technique, however, it may sometimes derive an incorrect map. Our new proposed technique overcomes this problem using recurrence relations. It first obtains the run history. Secondly, it performs regression analysis on loop iteration and variables based on the run history, and finally it derives a recurrence relation on the loop and variables occurring in the loop body. Experiments confirms that it can derive useful maps that we cannot derive by the existing technique.

Keywords: loop, static analysis, recurrence relation, run history, map

1 INTRODUCTION

Program analysis methods are divided into two categories: (1) Static Program Analysis and (2) Dynamic Program Analysis. Static program analysis methods do not execute the target program while dynamic program analysis methods execute the target program.

Static program analysis methods use many concrete methods such as symbolic execution [1] and model checking [2]. Recently yet another approaches have emerged. For example, heuristic methods [3], automatic predicate abstraction based methods, such as SLAM [4], BLAST [5] and so on. Static program analysis methods using logic sometimes utilize SAT/SMT solvers [6]. SAT/SMT solvers are enhanced SAT solvers with background theories.

For dynamic program analysis, Daikon [7] is a famous tool. It derives program assertions from data obtained from execution logs of the target programs. The execution is usually performed many times in order to infer accurate assertions. Recently approaches based on regression analysis [8] have been proposed [9]. Le [9] proposed a method that derives a map between variables before and after a target loop structure using regression analysis. It firstly executes the target loop many times with varying input values. Based on the data obtained by the execution, it then perform regression analysis. The analysis infers a map between variables before and after a target loop. Therefore, it easily analyzes program with loop structure. It, however, can deal with only linear and quadratic maps. Consequently, it cannot infer an exponent map which represents Fibonacci sequences.

Our proposed approach overcomes this problem by combining the general term and the result of the regression analysis.

The rest of this paper organized as follows. Section 2 provides disadvantages of the existing methods as preliminaries. Section 3 gives the proposed method. Sections 4 and 5 show experimental results and discussion, respectively. Finally, Section 6 summarizes this paper.

2 PRELIMINARIES

In general, static analysis approach sometimes suffers loop structures. Such techniques are not omnipotent due to memory limitation and calculating complexity. For this reason, some of the existing methods contrive several methods, such as bounded unwinding [10], user-specified time-out mechanism [11], and so on. In [12], S2E utilize Path Selection function, which enables us to control the termination with multiple criteria. For example, PathKiller can stop loop iteration up to a user specified number. As another approach, Xie *et al.* [13] proposed a method, which returns an *unknown* value for a variable that cannot be analyzed. Le [9] proposed a method based on regression analysis. It can derive a map between variables before and after a given loop structure. It claims that the method can derive more accurate map than others.

The approach in [9] uses approximation functions in Table 1.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \beta_4 x_1^2 + \beta_5 x_2^2$$

(x_1 and x_2 the input and the output, respectively)

There are many cases that the functions in Table 1 are not applicable. For example, a program for Fibonacci numbers in Listing 1 cannot be applied the functions in Table 1.

Listing 1: Program for Fibonacci Numbers

```
public class TestFibo {
    public int cf(int n) {
        int current = 0;
```

Table 1: Transformation Table for loop

Model	example	values for $\beta_0 \sim \beta_5$
Constant	$y = 0$	0
Simple Linear	$y = x_1$	0 except for β_1
Multiple Linear	$y = 2 \cdot x_1 + x_2$	$\beta_4 = 0, \beta_5 = 0$
Polynomial Linear	$a = x_1^2 + x_1 \cdot x_2$ if $x_1 > 0$	
Piece-wise Linear	then $y = x_1$ else $y = 3$	cannot be expressed

```

int prev = 1;
int prevprev = 0;
if (n > 0){
    for (int i = 0; i < n; i++) {
        current = prev + prevprev;
        System.out.print(current + " ");
        prevprev = prev;
        prev = current;
    }
    return current;
} else {
    System.err.println(
        "Input is less than 1");
    return -1;
}
}
}

```

3 OUR PROPOSED METHOD

Hereafter, we refer the approach in [9] as the existing method. First, we describe a difference between the existing method and our proposed method.

Our proposed method derives a map between variables before and after the target loop. The approach easily deals with loop structures.

The existing method analyses the map via regression analysis only. It, however, loose precision of the approximation. Our proposed method uses regression analysis only to derive a map between arguments (of the given Java method) and the number of iteration of the target loop (of the given Java method). A relation between n , the number of iteration of the target loop, and variables of the loop, is obtained by static analysis and an analysis tool for mathematics like Mathematica. This combination produces high accurate approximation.

3.1 The Outline of Our Proposed Method

Here, we give outline of our proposed method.

The inputs and the outputs of our method is summarized as follows.

- input: a Java method with a single loop structure
- output: an approximation of the loop structure in the method, if success. Otherwise failure.

Note that for a method with multiple loop structure, we can divide the method into several methods where each of them has a single loop structure. Thus, for a method with multiple loop structure, our proposed method is also applicable. For nest loops, it is known that such loops can be translated into a single loop. Therefore, in principle, such a method is also applicable.

We limit the class of input Java method as follows because we use SAT/SMT solvers at after analysis stages. Types allowed are bool, byte, short, int, float, double, and arrays of them. The proposed method cannot deal with String. For control structures, it allows to use `if`, `for`, `while` sentences.

Through the section, we use the following variables for their specific purposes.

x : the arguments of the given (target) method.

y : the variables which a user want to analyze.

n : the number of times of iteration of the loop.

Figure 1 shows the architecture of a proto-typed tool of our proposed method.

The procedure is summarized as follows.

Step 1: It adds properly print statements to the target source code in order to store execution logs on y and others.

Step 2: It executes the program with varying x and obtains enough execution logs.

Step 3: It analyzes the logs and obtains the execution paths, relations between x and the execution paths, and a record on n .

Step 4: Using regression analysis, it infers a relation between x and n .

Step 5: As a recurrence relation, it obtains a relation between y_s at the entry point and y_e exit point of the loop body.

Step 6: It solves the close-form of the recurrence relation obtained at **Step 5**.

Step 7: It finally calculates an expression representing a relation between y , x , and n by integrating **Step 3**, **Step 4**, and **Step 6**.

In the following subsections we will explain **Step 1**, **3**, **4**, **5**, **6**, and **7** which are important steps of our proposed method.

3.2 Step 1

In order to store execution logs, we add print sentences to the target source code. This step is similar to the Instrument step of Daikon [7], a famous tool for detecting invariants of programs.

JDT [14] is used to implement such instrument.

3.3 Step 3-1

Many paths are considered with regard to the conditions in the loop structure. Thus, we have to enumerate every pattern of paths. We call any path enumerated an execution path (in the loop). In general, the number of if statements is i , then there are 2^i execution paths at most.

3.4 Step 3-2

Here, we obtain a relation between x of the given method and sentences in the loop body.

The execution logs contains records on x and the number of occurrences of the execution path.

We explain more precisely using an example in Figure 2.

Let us assume that left upper code is the target method. The method contains two execution paths as shown in Figure 2.

We can see from Figure 2, that if the argument i is 25, then Execution Path2 (EXP2) occurs twice and EXP1 occurs three times.

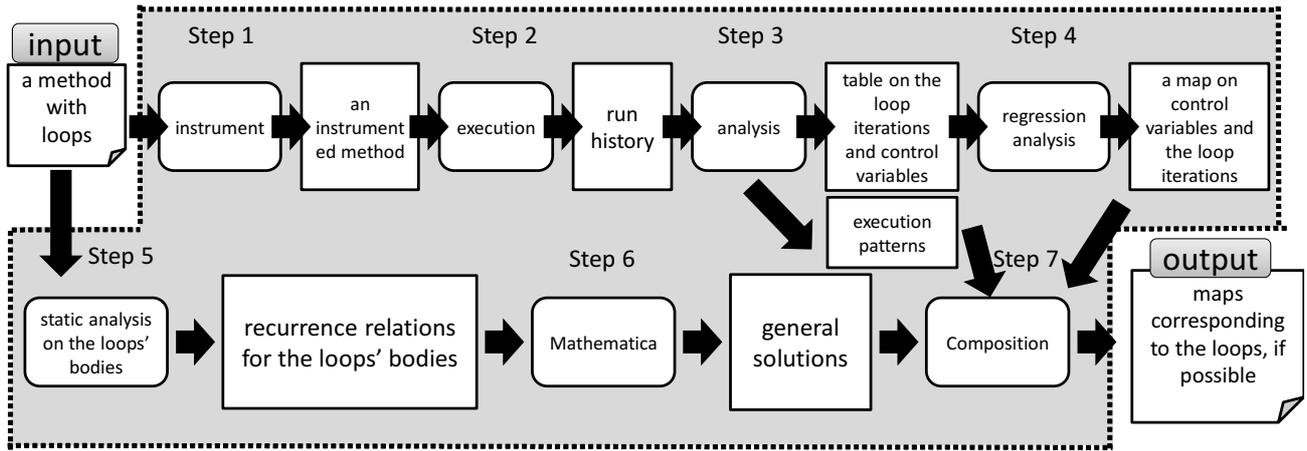


Figure 1: Tool overview

Figure 2 shows only three tuples, however, in actual, we obtain these tuples with more than 100 times execution varying the values of x . The same value 100 is used in the existing method. Many studies including [15] have proposed how to generate efficient values of the inputs (arguments).

3.5 Step 3-3

Next we confirm the order of the executing paths. Figure 2 shows the situation that first EXP2 is executed twice and then EXP1 is executed three times.

Therefore, we can represent this sequence as the following regular expression.

$$(EXP2)^*(EXP1)^* \tag{1}$$

At **Step 3-3**, we obtain such regular expression on the execution patterns.

Recall Figure 2. We can observe that the sequence is a sequence where first EXP2 is executed twice and then EXP1 is executed three times. We call such a pattern an execution pattern.

We enumerate every execution pattern from the execution logs. For each execution pattern, we abstract the constants representing the number of occurrence with Kleene closure symbol *. For example, three times occurrence of execution path EXP1 is abstracted into $(EXP1)^*$.

For a loop, we can obtain a set execution patterns.

For simply, hereafter we consider execution patterns in a form of $(EXP1)^*(EXP2)^* \dots (EXPn)^*(n > 0)$. For other cases, we return failure of analysis.

3.6 Step 4

Here, we describe how to derive a relation between x and the number of occurrence of an execution path (which is obtained at **Step 3-2**).

We use R[16], a regression analyzer.

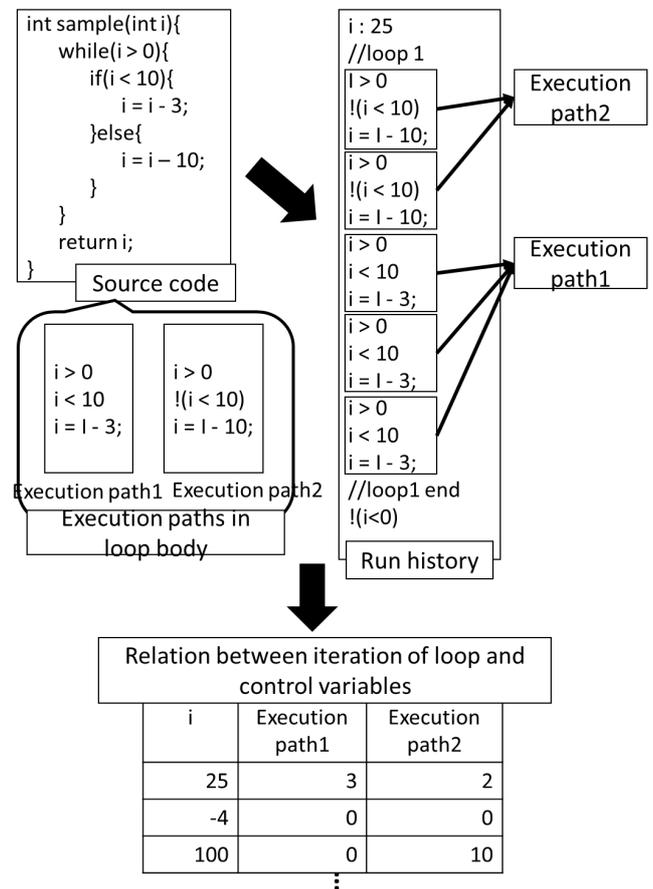


Figure 2: Analysis of Run history

Figure 3 shows relations between arguments (i and j) and an execution path named “loop.” Loop0.0 stands for “loop.” Variable i and j are integers.

A plot locating in the first row and the second column in Figure 3 shows a relation between i and j .

In a similar way, plots locate in the first row and the third

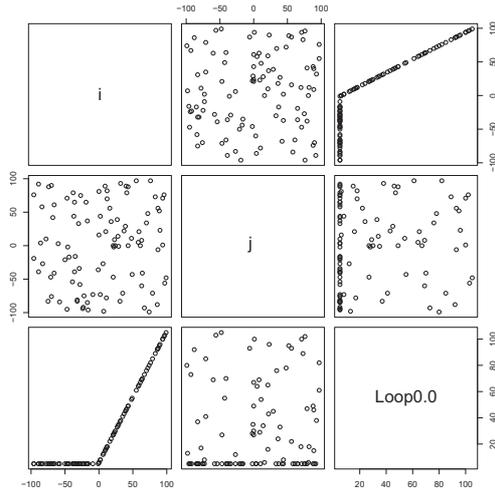


Figure 3: Relation between loop iteration and control variables

column, and in the second row and the third column show relations between i and the number of times of execution of “loop,” and between j and the number of times of execution of “loop,” respectively.

The plot locating in the first row and the second column in Figure 3 indicates that there is no correlation between i and j due to their randomness.

Additionally, we can find that there is no correlation between j and the number times of execution. However, there is large correlation between i and the number times of execution. For a case that i is negative, the number of times of execution of “loop” becomes 0. For a case that $i > 0$ holds, the number of times of execution of “loop” becomes i .

Such a relation can be obtained using regression analysis for each execution path.

For example, let us assume that arguments be i and j .

Let $a_0, a_1, a_2 \dots$ be coefficients.

The following model (expression) can be used in regression analysis.

$$n = a_0 + a_1i + a_2j + a_3ij + a_4i^2 + a_5j^2 \quad (2)$$

For Figure 3, we obtain a result that a_1 equals 1 and other coefficients are 0. Thus, we obtain a relation $n = i$.

Note that n is the number of times of iteration. Thus, it does not have a negative value. We assume that $n = 0$ when $n < 0$ for later analysis steps.

For the case of failure of regression analysis, we return analysis failure.

3.7 Step 5

For each execution path, we derive a relation between variables y_s and y_e .

First, a series of assignment statements of the execution path into SSA (static Single Assignment) form [17]. In SSA form every variable allows at most once assignment. Therefore, an original variable is, in general, divided into several

variables when it is involved in several assignment statements. For such a case, the divided variables are distinguished by its own suffix.

Listing 2 and Figure 4 show an example of an SSA form.

An execution path in Listing 2is translated into an SSA form shown in Figure 4. The execution path uses variable z and y . Their corresponding variables for the first assignment, are represented as z_0 and y_0 . At line 1, $z + y$ is assigned for z . In such a case, variables z_0 and z_{+1} are used. In a similar way, at line 3, z_2 is used. The suffixes play roles to distinguish variable z at different locations.

Listing 2: SSA example original Java

```
z=z+y;
y=y+1;
z=z+1;
y=z+y;
```

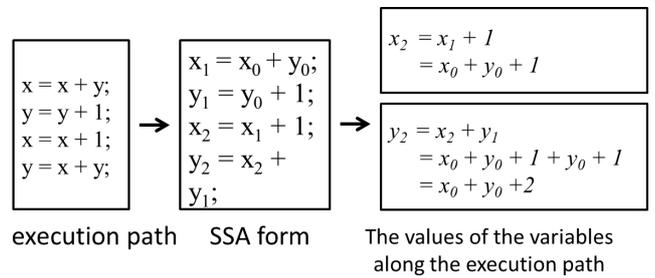


Figure 4: SSA form

Next using the SSA form, we derive a recurrence relation on the variables.

In Figure 4, the first values of the variable z and y are represented as z_0 and y_0 . The final values are represented as variables z_2 and y_2 . Using the SSA form, we can infer that z_2 equals $z_1 + 1$ and also z_1 equals $z_0 + y_0$. Thus, z_2 and y_2 equal $z_0 + y_0 + 1$ and $z_2 + y_1$, respectively. Because y_1 equals $y_0 + 1$, we infer that y_2 equals $z_0 + y_0 + 2$.

The obtained equations can be represented as recurrence relations. We then obtain the following recurrence relation.

$$z[n + 1] = z[n] + y[n] + 1, \quad y[n + 1] = z[n] + y[n] + 2 \quad (3)$$

Here, $z[0]$ and $y[0]$ stand for the seed values, i.e., z_0 and y_0 for the variables z and y . Variables $z[n]$ and $y[n]$ stand for the values of z_n and y_n that are the values at the point where the n times iteration of the loop.

3.8 Step 6

Here, we solve the recurrence relation.

For example, we can obtain the following recurrence relation from an SSA form in Figure 4.

$$z[n + 1] = z[n] + y[n] + 1, \quad y[n + 1] = z[n] + y[n] + 2 \quad (4)$$

We use Mathematica [18] for solving recurrence relations.

Let us assume that the seed values of z and y are z_0 and y_0 , respectively. We can obtain a closed-form solution for the recurrence relation using Mathematica, as follows.

$$z[n] = \frac{1}{2}(-4 + 3 \cdot 2^n + 2^n y_0) \quad (5)$$

$$y[n] = \frac{1}{2}(-2 + 3 \cdot 2^n + 2^n z_0) \quad (6)$$

It is hard to obtain such a complex expression by the existing method [9].

3.9 Step 7

Here we integrate the obtained analysis at the previous steps and obtain the final map.

Let us assume that the execution pattern is $(EXP1)^*(EXP2)^* \dots (EXPk)^*$, and that the number of the execution times of $EXPi$ is m_i .

\mathbf{y}_0^i stands for the initial values at the entry of $EXPi$

Let $\mathbf{y} = F_i(\mathbf{y}_0^i, n)$ be the map obtained at **Step 6**. Let G_i be the map obtained at **Step 5**, where $m_i = G_i(\mathbf{x})$.

1. $k = 1$ holds
2. $k > 1$ and $\forall i \exists c : 0 < i \leq k, (G_i = c\mathbf{x} \text{ or } G_i = c)$ holds
3. $\forall i, j : 0 < i, j \leq k, G_i = G_j$ holds
4. otherwise

For case (1), the final map is $\mathbf{y} = F_1(\mathbf{y}_0^1, G_1(\mathbf{x}))$.

For cases (2) and (3), the final map is $\mathbf{y} = F_n(\dots F_2(F_1(\mathbf{y}_0^1, G_1(\mathbf{x})), G_2(\mathbf{x})), \dots, G_n(\mathbf{x}))$.

For the case (4), we conclude that it is failure analysis.

We explain the case (2):

Let assume $(EXP1)^*(EXP2)^*$ be the execution pattern.

Let us consider a situation where when $EXP1$ is executed n times, then the value of variable y increases by n , and if $EXP2$ is executed n times, then the value of variable y increases by $10n$.

In such a case, equations $F_1(y_0^1, n) = y_0^1 + n$ and $F_2(y_0^2, n) = y_0^2 + 10n$ holds. Additionally let us assume that when $i > 0$ $EXP1$ and $EXP2$ are executed i and 1 times, respectively; and that when $i \leq 0$ $EXP1$ and $EXP2$ are not executed.

By integrating the above all, we can obtain the final map $y = y_0^1$ for $x \leq 0$, and $y = y_0^1 + x + 10$ for $x > 0$.

4 EXPERIMENTS

The setup of the experiments is summarized as follows.

- OS: Windows 7 Enterprise 64bit
- CPU: Intel Xeon E5-2609 2.40GHz \times 2
- Memory: 48.0GB
- Java: JRE7
- R: version 3.0.2
- Mathematica 9.0.0
- Z3: z3-4.3.0

4.1 The overview of the experiment

The research questions are summarized as follows.

RQ1 Map Quality: Is the obtained map accurate?

RQ2 Map Novelty: Is the obtained map derived from the existing method?

Table 2: Target Programs

Program	LOC	#of loop st	# of if st
Fibonacci	20	1	1
Newton	30	1	1
Draw pictures	39	1	1
Draw parabola	77	3	10
Summation	20	1	1
Power series	20	1	1

Table 2 summarizes the target programs. Each program has loop structures.

4.2 The Results

Tables 3 and 4 show execution times and results.

5 DISCUSSION

5.1 RQ1

With Mathematica, we can correctly derive closed-form solutions of the recurrence formulae obtained from the programs. For some of other programs not shown in Table 2, we cannot derive their closed-form solutions. The reasons are that (1) Mathematica cannot deal with them, and (2) in general, every recurrence relation does not has to have its closed-form solution exists. For such a case, existing methods also cannot be applied.

5.2 RQ2

The existing method, in principle, cannot derive a correct map for Fibonacci. The existing method approximates it as quadratic equations. It, however, has large relative errors for a large input. Thus, advantage of our proposed method is confirmed.

5.3 Execution times

Most of execution times are occupied by Mathematica. Particularly, solving a closed-form solutions take many costs.

6 CONCLUSION

This paper proposed a new method to infer a map between variables before and after a given loop structure. The experimental results show that our proposed method can derive complex map which existing methods cannot derive.

Future work include to apply concolic testing on our derived maps, in order to perform efficient and further analysis.

Table 3: Execution Times in Each Step (sec.)

Program	1	2	3	4	5	6 and 7	total
Fibonacci	2.6	0.62	0.1	4.4	4.4	0.72	8.4
Newton	2.3	0.36	0.0	2.4	1.3	0.64	5.8
Draw Pic.	2.7	0.59	0.0	3.0	2.2	0.75	7.0
Draw Para.	2.5	1.1	0.0	2.9	2.9	0.65	7.2
Summation	2.4	0.57	0.05	2.4	2.4	0.66	6.0
Power	2.3	0.38	0.0	3.5	3.5	0.75	6.9

Table 4: Experimental Results

Program	the existing method	our method
Fibonacci	×	✓
Newton	×	✓
Draw Pic.	✓	✓
Draw Prara.	×	✓
Summation	✓	✓
Power	×	✓

Acknowledgment

The research is also being partially conducted as Grant-in-Aid for Scientific Research C (16K00094) and S(25220003).

REFERENCES

- [1] P. Godefroid, N. Klarlund, and K. Sen: "DART: directed automated random testing," ACM SIGPLAN Notices, vol.40, no.6, pp.213–223 (2005)
- [2] W. Visser, K. Havelund, G. Brat, S. Park, and F. Lerda: "Model checking programs," Automated Software Engineering, vol.10, no.2, pp.203–232 (2003)
- [3] P. Cousot: "Proving program invariance and termination by parametric abstraction lagrangian relaxation and semidefinite programming," Proceedings of the 6th International Conference of VMCAI 2005, Vol. 3385, pp.1–24, Lecture Notes in Computer Science (2005)?D
- [4] T. Ball and S.K. Rajamani: "The slam project: Debugging system software via static analysis," Proceedings of the 29th ACM SIGPLAN-SIGACT POPL'02, pp.1–3 (2002)
- [5] T.A. Henzinger, R. Jhala, R. Majumdar, G.C. Necula, G. Sutre, and W. Weimer: "Temporal-safety proofs for systems code," Proceedings of the 14th International Conference on Computer Aided Verification, CAV 2002, pp.526–538 (2002)
- [6] A. Biere, M. Heule, H. Van Maaren, and T. Walsh: "Handbook of Satisfiability," IOS press (2009)
- [7] M.D. Ernst, J. Cockrell, W.G. Griswold, and D. Notkin: "Dynamically discovering likely program invariants to support program evolution," IEEE TSE, vol.27, pp.1–25 (2001)
- [8] M. Younger: "Handbook for Linear Regression," Duxbury Resource Center (1979)
- [9] W. Le: "Segmented Symbolic Analysis," Proceedings of the 2013 International Conference on Software Engineering, pp.212–221 (2013)
- [10] D.R. Cok and J.R. Kiniry: "Esc/java2: Uniting esc/java and jml: Progress and issues in building and using esc/java2 and a report on a case study involving the use of esc/java2 to verify portions of an internet voting tally system," Proceedings of Construction and Analysis of Safe, Secure and Interoperable Smart Devices: International Workshop, CASSIS 2004, vol.3362, pp.108–128, Lecture Notes in Computer Science (2005)
- [11] C. Cadar, D. Dunbar, and D. Engler: "KLEE: Unassisted and Automatic Generation of High-coverage Tests for Complex Systems Programs," Proceeding of the 8th USENIX Conference on Operating Systems Design and Implementation, pp.209–224 (2008)
- [12] V. Chipounov, V. Kuznetsov, and G. Candea: "S2E: A Platform for In-vivo Multi-path Analysis of Software Systems," Proceedings of the 16th International Conference on Architectural Support for Programming Languages and Operating Systems, pp.265–278 (2011)
- [13] Y. Xie, A. Chou, and D. Engler: "ARCHER: Using Symbolic, Path-sensitive Analysis to Detect Memory Access Errors," Proceedings of the 9th European Software Engineering Conference Held Jointly with 11th ACM SIGSOFT International Symposium on Foundations of Software Engineering, pp.327–336 (2003)
- [14] "Eclipse Java development tools (JDT)," (accessed 2015-05-05). <http://www.eclipse.org/jdt/>.
- [15] K. Kobayashi, Y. Sasaki, K. Okano, and S. Kusumoto: "Automated assertion generation using PDF and SMT-Solver," IEICE Transaction on Information and Systems, JD, vol.96, no.11, pp.2657–2668 (2013) (In Japanese)
- [16] "the R statistical package," (accessed 2015-05-05). <http://cran.r-project.org/>.
- [17] E. Clarke, D. Kroening, and F. Lerda, "A Tool for Checking ANSI-C Programs," Proceedings of the 10th International conference on Tools and Algorithms for the Construction and Analysis of Systems, pp.168–176 (2004)
- [18] "Mathematica: Wolfram Research," (accessed 2015-05-05). <https://www.wolfram.com/>.
- [19] K. Sen, D. Marinov, and G. Agha: "CUTE: A Concolic Unit Testing Engine for C," SIGSOFT Software Engineering Notes, vol.30, no.5, pp.263–272 (2005)
- [20] R. Majumdar and K. Sen: "Hybrid Concolic Testing," Proceeding of the 29th International Conference on Software Engineering, pp.416–426 (2007)

A Method for verifying equivalence of functions in C language

Satoshi Harauchi^{*}, Kozo Okano^{**}, and Shinpei Ogata^{**}

^{*}Advanced Technology R&D Center, Mitsubishi Electric Corporation, Japan
Harauchi.Satoshi@bc.MitsubishiElectric.co.jp

^{**}Computer Science & Engineering, Shinshu University, Japan
{okano, ogata}@cs.shinshu-u.ac.jp

Abstract - Software programs are often revised when we develop or maintain them. Such revisions are performed in order to speed up their algorithms or to modify them into more flexible. Such modification is called ‘Refactoring’. When modifying programs, tests should be performed in order to satisfy their external specifications. We generally execute test cases for the modified programs as regression testing. In general, it needs time and efforts, and sometimes causes an error due to the omission of test cases. In this paper, we propose a method to check the equivalence of software programs between before and after changes. We focus on functions in programs. The proposed method utilizes software analysis workbench (SAW) which is an open source software. The proposed method consists of following procedures. The method generates checking scripts for SAW from original programs. We then apply SAW with the generated scripts to the revised programs. The result shows whether both programs are equivalent, and suggests the counter example when programs are not equivalent. This paper describes our method, application for C programs and evaluation of the method. The result of evaluation for several C programs shows that the method is effective.

Keywords: C language, Software analysis workbench

1 INTRODUCTION

Software programs are frequently changed or modified when we develop or maintain them. When we develop software framework, we often change framework in order to apply it to various systems. When we implement specific algorithm to the software, we occasionally modify the software to speed up the algorithms. Such changes or modifications are called ‘Refactoring’ [1]. Refactoring is a kind of changes to improve software while maintaining external specifications.

When we change or modify programs, tests should be executed so as to keep specifications. We perform unit tests as regression tests by using the existing test cases for modified software. Mostly it takes time and efforts, and sometimes leads to the occurrence of bugs due to the failure to cover every test case. Therefore it would be preferable to support such tests when software is modified.

In this paper, we propose a method to verify the equivalence of software programs between before and after modification. Proposed method focuses on functions in C language. The method achieves verification by the use of

software analysis workbench (SAW) [2][3] which is an open source software. The method shows whether both programs are valid or invalid, and moreover indicate counter example when detecting invalidity. The method aims at giving tools for verifying the equivalence of the programs and reducing the time and effort needed for tests.

The method is organized as follows. At first, the method produces LLVM [4] bitcode from C programs before and after modification. LLVM bitcode is a kind of intermediate expressions for programming language. The method then generates the script for verification from the both programs. The script contains names of modules, types of parameters and outputs described in the functions and instructions to verify functions. Next we execute SAW tool to verify programs with LLVM bitcodes and the script. SAW makes use of SAT solvers or SMT solvers [5][6] for formal verifications such as Z3 [7] and Yices2 [8], theorem provers such as abc [9]. SAW calls selected solver internally when running verification. The result of verification provides the equivalence between both programs.

We implement the method for C language and perform two evaluations. First evaluation focuses the capability to check the equivalence. We prepare several sample programs as original programs and create different programs corresponding to the programs. We then prepare the script for verification which covers all of possible inputs parameters. We utilize SAW with the script by checking all of outputs for inputs. The result of evaluation suggests that SAW has some constraints about parameters. Second one focuses the validity of generating the script. We check the generated script by confirming whether we obtain the same result as first evaluation. The generation turns out to be valid from the successful result of evaluation.

This paper shows the method in detail. First, we describe our method in section 2. We then describe implementation and evaluation of the method in section 3.

2 PROPOSED METHOD

Proposed method consists of three steps shown in figure 1. We describe the details of each step in this section.

2.1 STEP 1

The method generates LLVM bitcode from original function and modified function. LLVM code is an intermediate language used for C++, Objective-C in addition to C language. Therefore functions can be written in the different languages.

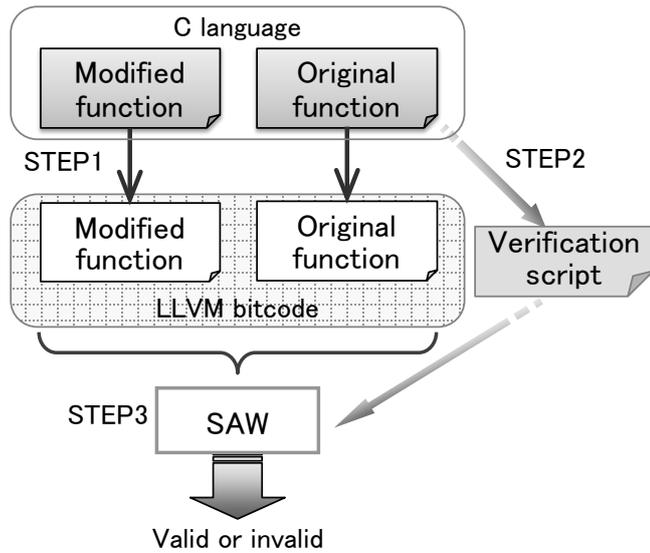


Figure 1: An overview of proposed method

```

1 load <- llvm_load_module "$1 .bc";
2 $2 <- fresh_symbolic "$3" { $4 };
3 $5 <- fresh_symbolic "$6" { $7 };
4
5 let allocs_org = [ $8 ];
6 let allocs_mod = [ $8 ];
7 let inputs_org = [ $9 ];
8 let inputs_mod = [ $9 ];
9 let outputs_org = [ $10 ];
10 let outputs_mod = [ $10 ];
11
12 print "let t1";
13 t1 <- llvm_symexec load "$11 allocs_org
14 inputs_org outputs_org true;
15 print "let t2";
16 t2 <- llvm_symexec load "$12 allocs_mod
17 inputs_mod outputs_mod true;
18 print "t1 = t2";
19 thm <- abstract_symbolic {{ t1 == t2 }};
20 r <- prove abc thm;
21 print r;
  
```

Figure 2: An example of the template

Table 1: The meaning of symbols for figure 2

\$1	file name of bitcode
\$2	any name of the variable used in the script for input parameter
\$3	a name of input parameters for specified function
\$4	size of input parameter type and array size of \$3
\$5	any name of the variable used in the script for output parameter
\$6	a name of output parameters for specified function
\$7	size of output parameter type and array size of \$3
\$8	array size of parameter
\$9	name of input parameter and array size of parameter
\$10	name of output parameter and array size of parameter

\$11	name of original function
\$12	name of modified function

2.2 STEP 2

This step generates the script for verifying the equivalence of functions. SAW allows users to execute any verifications by the use of symbols. SAW processes the verification with the description called AIG (And-Inverter Graphs) [10].

The script follows the rule defined by SAW. The method prepares several templates of script. The template is determined according to the type of parameters in functions. Only integers are available for parameter due to the SAS constraints.

Figure 2 shows an example of the templates for the parameter including array. The template has variables from \$1 to \$12 whose meanings are shown in table 1. From \$1 to \$3, \$5, \$6, \$11 and \$12 are required in common. \$4, \$7 and From \$8 to \$10 are needed for the parameter with array. \$4 and \$7 declares that variable \$2 and \$5 is the array with specified size. \$8 defines the allocation of the array. \$9 and \$10 describes the name of parameters and the size of array by the use of variable \$2 and \$5.

The method parses original function and modified function written in C language and extract parameters in functions. The method then select an appropriate template according to the extracted parameters and replace symbols in the template to the actual string in functions.

2.3 STEP 3

The method verifies functions by the use of SAW. The method uses LLVM bytecode and the generated script. SAW uses SAT/SMT solvers in verification. The name of solvers should be described in the script. Both functions are proved to be equivalent when the method provides the successful information. The method shows invalidity with counter example when functions are different.

3 EVALUATION

3.1 The capability to check the equivalence

We prepare several programs in order to check the capability for the verification. We create three kinds of software as original programs. We then develop another programs with the changes of the algorithm while keeping specifications. Next we create the script and apply SAW to created programs and the script. We use solvers for abc, yices, cvc4 [10] and z3 and measure the time spent for verifying with each solver. We evaluate the capability by checking whether SAW provides the validity or not.

Computer specification used for evaluation is described as follows.

- OS: OS X El Capitan version 10.11.6
- CPU: Intel Core i7-4770HQ/2.2GHz
- Memory: 16.00GB
- LLVM Clang: version 3.8.1

• SAW: version 0.2

3.1.1. Sort program

This section evaluate sort programs. We prepare bubble sort program as an original program. We create quick sort program as a modified program. The original programs and quick sort programs are shown in figure 3 and figure 4, respectively. The script is shown in figure 5.

```

1 void BubbleSort( int *input , int *output , int size)
2 {
3     int i, j, tmp;
4     for ( i=0; i<size ; i++)
5     {
6         output [i] = input [i];
7     }
8     tmp = 0 ;
9     for ( i=size ; i>0; i-- )
10    {
11        for ( j=1; j<i ; j++)
12        {
13            if ( output [j-1] > output [j] )
14            {
15                tmp = output [j-1];
16                output [j-1] = output [j];
17                output [j] = tmp;
18            }
19        }
20    }
21 }

```

Figure 3: Bubble sort program

```

1 void quick ( int *list, int left, int right)
2 {
3     int i , j,pivot;
4     int tmp;
5     tmp = 0;
6     i = left;
7     j = right;
8     pivot = list [ ( i+j ) / 2 ];
9     while ( 1 )
10    {
11        while ( list [i] < pivot ) i++;
12        while ( pivot < list [j] ) j--;
13        if ( i >= j ) break ;
14
15        tmp = list [i] ;
16        list [i] = list [j] ;
17        list [j] = tmp ;
18        i++;
19        j--;
20    }
21    if ( left < i-1 ) quick ( list , left, i-1 );
22    if ( right > j+1 ) quick ( list , j+1, right);
23 }
24
25 void QuickSort( int *input , int *output , int size)
26 {

```

```

27     int i;
28     for ( i =0; i<size ; i++)
29     {
30         output [ i ] = input [ i ] ;
31     }
32     quick ( output , 0 , size-1 );
33 }

```

Figure 4: Quick sort program

```

1 m <- llvm_load_module "sort_test.bc";
2 xs <- fresh_symbolic "xs" { |[3] [8] | };
3 outs <- fresh_symbolic "outs" { |[3] [8] | };
4 let num = 3;
5 let ctr = { { 3 : [ 8 ] } };
6 let allocs = [ ( "x" , num) , ( "out" , num) ] ;
7 let inputs = [ ( "*"x" , xs , num)
8               , ( "*"out" , outs , num)
9               , ( "size" , ctr , 1)
10              ];
11 let outputs = [ ( "*"out" , num)
12               ];
13
14 print "x [3] [8]" ;
15 print "let t1";
16 t1 <- llvm_symexec m "BubbleSort" allocs inputs
17   outputs true;
18 print "let t2";
19 t2 <- llvm_symexec m "QuickSort" allocs inputs
20   outputs true;
21 print "t1 = t2";
22 thm1 <- abstract_symbolic { { t1 == t2 } };
23 prove_print abc thm1;

```

Figure 5: Verification script

The result of evaluation is show in table 2 and figure 6. N/A in table 2 means the failure of measurement due to long process. The result shows that time of verification grows exponentially with the array size. This may be caused by the existence of recursive functions in quick sort programs. Recursive calls may be difficult to be verified.

The result also shows the solver cvc4 and 8 bit size are not compatible because the time of 8 bit size is longer than that of 32 bit size. This problem can be solved by using solver abc instead of cvc4 in case of verification for 8 bit size.

Table 2: The time of verification for sort programs

parameter size	array size	abc(s)	yices(s)	cvc4(s)	z3(s)
8 bit size	2	0.622	0.679	0.693	0.670
	3	1.880	1.876	45.210	1.861
	4	16.079	16.280	N/A	15.829
	5	N/A	N/A	N/A	N/A
32 bit size	2	0.802	0.893	0.863	0.801
	3	5.362	5.455	5.559	5.607
	4	76.369	72.178	75.158	76.614
	5	N/A	N/A	N/A	N/A

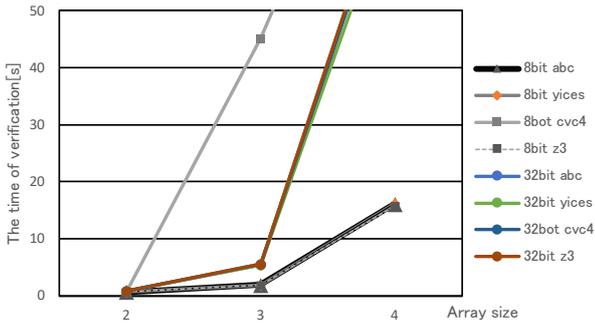


Figure 6: The time of verification for sort programs

3.1.2. Gaussian filter programs

In general, some methods exist for accelerate software programs. This section evaluate the programs using MapReduce which is one of the accelerating technique. MapReduce divides original data into several distributed data and processes the data in parallel. Each data are processed on different computers through networks and the results of processing are aggregated into the main computer.

```

1 #define M_ROW (5) /* row */
2 #define M_COL (5) /* column */
3 void Gaussian(int input[M_ROW][M_COL] ,
4               int output[M_ROW][M_COL])
5 {
6     int i,j,r;
7     for (i =0; i<M_ROW; i++){
8         for (j =0; j<M_COL; j++){
9             output [ i ] [ j ] = 0 ;
10        }
11    }
12    ...

```

Figure 7: Gaussian filter programs

Table 3: The time for gaussian filter programs

array size	abc(s)	yices(s)	cvc4(s)	z3(s)
1	0.606	0.523	0.535	0.522
2	0.773	0.806	0.802	0.761
3	1.467	1.309	1.347	1.229
4	2.776	2.475	2.335	2.471
5	4.925	4.785	4.76	4.283
6	8.148	7.756	7.69	7.863
7	12.913	12.541	12.669	12.962
8	20.286	21.073	19.598	19.868
9	31.353	33.305	30.247	30.515
10	45.55	47.312	44.401	45.123
11	71.122	65.47	67.11	64.215
12	90.89	93.246	89.43	89.433
13	133.911	121.653	119.532	123.861
14	169.562	167.409	159.992	164.164
15	216.018	210.152	209.849	209.157
16	272.676	279.172	277.165	285.011
17	367.315	350.251	346.219	377.539
18	452.035	442.479	443.856	435.871
19	558.74	540.154	550.934	547.763
20	683.641	674.93	654.051	657.453

In this section we evaluate gaussian filter programs as an application of pseudo MapReduce. We prepare filter program an original program. We create another filter program to which MapReduce is applied. We implement both programs with two-dimensional array and measure the time of verifications. Figure 7 shows a part of the original programs.

The result of evaluation is show in table 3 and figure 8. The result shows that verification is successful even for the array size 20 although sort programs fail to verify for only the array size 4. This may be because gaussian filter programs are simple compared to sort programs. Another result shows that the time verification has almost no difference between solvers.

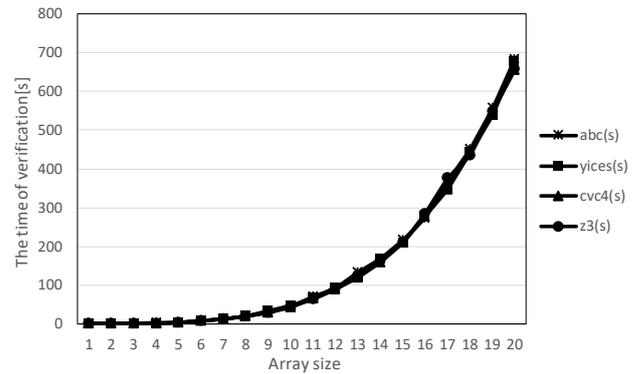


Figure 8: The time for gaussian filter programs

3.1.3. Binary tree search programs

This section evaluate binary tree search programs. We execute the verification of equivalence between search algorithm. We prepare pre-order searching program as an original program. We create in-order searching program and post-order searching program as modified programs. Figure 9 shows a part of the original programs. As shown in figure 9, the parameter of functions contain structure.

```

1 struct BTREE {
2     int value ;
3     struct BTREE* left;
4     struct BTREE* right;
5 };
6
7 Btree_t auto_create_binary_tree (int *list)
8 {
9     int i;
10    Btree_t node [NODE_NUM] ;
11    for (i=0; i<NODE_NUM; i++){
12        node [i] = malloc( sizeof( struct BTREE ));
13        node [i]->value = i;
14    }
15    ...

```

Figure 9: Binary tree search program

Unfortunately we cannot measure the time of verification. We obtain the result of invalidity though searching programs are equivalent. The result of investigation shows that type of structure causes invalidity. SAW cannot manage

structure and access each value of members in structure. The result of evaluation suggests that SAW has constraints about types of parameters.

3.2 Generating script

We implement the tool for generating script and evaluate the validity of tools. We check the generated script by comparing with the script created in previous evaluation. We evaluate the script for sort programs in section 3.1.2 and gaussian filter programs in section 3.1.3.

Figure 10 shows the script generated from figure 2 for gaussian filter programs shown in figure 7. Line 2 and 3 in figure 10 declare the variable with two-dimensional array with the array size 5 and 32 bit parameter. Line 5 and 6 allocate both input and output parameter. From line 7 to 10 describe the name of parameters and the size of parameter.

We evaluate the verification using the generated script. We obtain the same result of verification as in previous evaluation.

```

1 load <- llvm_load_module "gauss.bc";
2 input <- fresh_symbolic "input" { |[ 5 ] [ 5 ] [ 3
  2 ] | };
3 output <- fresh_symbolic "output" { |[ 5 ] [ 5 ] [ 3
  2 ] | };
4
5 let allocs_org = [ ("input", 5), ("output", 5) ];
6 let allocs_mod = [ ("input", 5), ("output", 5) ];
7 let inputs_mod = [ ("*input", input, 5), ("*
  output", output, 5)];
8 let inputs_org = [ ("*input", input, 5), ("*
  output", output, 5)];
9 let outputs_org = [ ("*output", 5) ];
10 let outputs_mod = [ ("*output", 5) ];
11
12 print "let t1";
13 t1 <- llvm_symexec load "GaussianOrg"
  allocs_org inputs_org outputs_org true;
14 print "let t2";
15 t2 <- llvm_symexec load "GaussianMod"
  allocs_mod inputs_mod outputs_mod true;
16 print "t1 = t2";
17 thm <- abstract_symbolic {{ t1 == t2 }};
18 r <- prove abc thm;
19 print r;

```

Figure 10: The generated script

4 CONCLUSION

This paper proposed a method to verify the equivalence of software programs between before and after modification on functions in C language. Our method suggests both programs are equivalent by verifying with software analysis workbench, and shows counter example in invalidity. This verification provides the reduction of the time and effort spent for comprehensive tests after modification.

We implement the method and perform two evaluations. In the first evaluation, we measured the time spent for executing verification for prepared programs. The result of

evaluation indicates that SAW has constraints for parameters defined in functions. In the second one, we check the validity of tools for generating the script. We confirm the equivalence between generated script and the script which provides successful result for first evaluation. Future work will apply our method to actual software projects and various kinds of functions used for products.

REFERENCES

- [1] Marin Fowler. Refactoring – Improving the Design of Existing Code. Addison-Wesley, 1999.
- [2] Galois, Inc.: “The Software Analysis Workbench,” <http://saw.galois.com/index.html/> <referred on 24th May> (2016)
- [3] R. Dockins, A. Foltzer, J. Hendrix, B. Huffman, D. McNamee, and A. Tomb: “Constructing Semantic Models of Programs with the Software Analysis Workbench,” In Proceedings of VSTTE 2016 (2016).
- [4] “LLVM”. <http://llvm.org/> <referred on 24th May 2016>
- [5] C. Barrett, A. Stump and C. Tinelli, “The SMT-LIB Standard Version 2.0,” (2010)
- [6] A. Cimatti, A. Griggio, B. Schaafsma, and R. Sebastiani: “The MathSAT5 SMT Solver,” In Proceedings of TACAS 2013, LNCS Vol. 7795, pp.93-107 (2013)
- [7] L. de Moura and N. Bjørner: “Z3: An efficient SMT solver,” In Proceedings of TACAS 2008, LNCS Vol. 4963, pp.337-340 (2008)
- [8] B. Dutertre: “Yices 2.2,” In Proceedings of CAV2014, LNCS Vol.8559, pp.737-744 (2014)
- [9] R. Brayton and A. Mishchenko: “ABC: An Academic Industrial-Strength Verification Tool,” LNCS Vol.6174, pp. 24-40 (2010)
- [10] C. Barrett, C. L. Conway, M. Deters, L. Hadarean, D. Jovanović, T.King, A. Reynolds, and C. Tinelli: “CVC4,” Proceedings of the 23rd international conference on Computer aided verification (CAV’11) pp. 171-177 (2011)
- [11] A. Darringer, W. H. Joyner, Jr., C. L. Berman, and L. Trevillyan: “Logic synthesis through local transformations,” IBM Journal of Research and Development. Vol. 25 (4): pp. 272-280 (1981)

A Rule-based Method of Stepwise Evaluating Class Diagrams

Shinpei Ogata[†], Kazune Miyajima[‡], Mizue Kayama[†] and Kozo Okano[†]

[†]Academic Assembly, Shinshu University, Japan

[‡]Graduate School of Science and Technology, Shinshu University, Japan

{ogata, kayama, okano}@cs.shinshu-u.ac.jp

Abstract - UML class diagrams constitute an essential specification in object-oriented development. The appropriate form of class diagrams in general cannot be uniquely determined because of the hierarchy and ambiguity of natural language terms. In the context of education, this implies that appropriate and efficient evaluation of many class diagrams is difficult for instructors. Thus, various methods of evaluating class diagrams that enhance the efficiency and appropriateness of the evaluation have been proposed. Appropriateness should precede efficiency in importance. However, thus far no adequate evaluation method exists. This paper therefore presents a rule-based method of stepwise evaluating class diagrams in order to address this problem. The effectiveness of the proposed method was evaluated by applying it to many class diagrams obtained from lectures given at Shinshu University.

Keywords: Class Diagram, Evaluation Support, Object-Oriented Development Education, Unified Modeling Language.

1 INTRODUCTION

UML class diagrams constitute one of the essential specifications in object-oriented development. A class diagram in general is a semi-formal artifact, because a part of the class diagram is informally expressed in natural language. Thus, the appropriate form of class diagrams that satisfies certain requirements frequently cannot be uniquely determined because of the hierarchy and ambiguity of natural language terms. In the context of education, this implies that instructors cannot easily evaluate a large number of class diagrams appropriately and efficiently.

On this background, various methods of evaluating class diagrams have been proposed that are aimed to enhance the efficiency and appropriateness of the evaluation [1]–[4], [6]–[8], [10], [12]–[15]. While some methods can efficiently evaluate learners' class diagrams by automatically comparing them with a corresponding reference diagram [1], [2], [8], the risk exists that they may decrease the appropriateness of the evaluation. For instance, when a class diagram is appropriate but not very similar to the reference model, the diagram may be evaluated inappropriately. To make matters worse, the users of the methods might not notice this situation. Appropriateness should precede efficiency in importance; however, no adequate evaluation method yet exists.

This paper therefore presents a rule-based method of stepwise evaluating class diagrams in order to resolve this problem. In our method, the rules for evaluating class diagrams are manually created by instructors, but by using our tool the templates of the rules can be automatically generated from

learners' diagrams. In other words, the instructors take responsibility for evaluating the class diagrams while the tool plays the role of aggregating the class diagrams' information for evaluation. Thus, this research focused on the effectiveness of the aggregation for manually evaluating a large number of class diagrams in the context of education.

The effectiveness of the proposed method was evaluated by applying it to class diagrams obtained from lectures given at Shinshu University.

2 Related Work

Many methods for supporting class diagram evaluation have been proposed: automatic evaluation using reference diagrams [1], [2], [8], evaluation of layouts from the viewpoint of architecture [3], explanations of modeling principles with modeling examples [7], automatic checking of syntactic and pragmatic errors and inconsistencies between diagrams [13], and check lists [15].

Some of these existing methods cannot automatically process the natural language used in the informal part of the class diagrams, e.g., class name, and attribute name. Meanwhile, automatic evaluation methods [1], [2], [8] offer efficient class diagram evaluation. In particular, UMLGrader [8] processes natural language automatically to some extent in order to cover the differences between the learners' diagrams and a corresponding reference diagram in terms of natural language expressions. However, automation that includes natural language processing frequently leads to error-prone evaluation, which is a critical risk in the context of education.

An instructor should take the responsibility for the education of his/her students, and thus should evaluate all their answers. Hence, it is important that he/she be able easily comprehend all the answers and evaluate them in detail. In this paper, we propose a method that facilitates efficient and appropriate class diagram evaluation in order to support the instructor in this activity.

3 Rule-based Method and Support Tool for Stepwise Evaluation of Class Diagrams

Fig. 1 shows the overview of our proposed tool. An instructor as a user interacts with the tool in order to evaluate class diagrams. Whereas all the actions provided by the tool are fully automated, all the instructor's actions are performed manually.

Step 1. The instructor inputs the class diagrams to the tool. Each of the class diagrams is created in a file of Astah

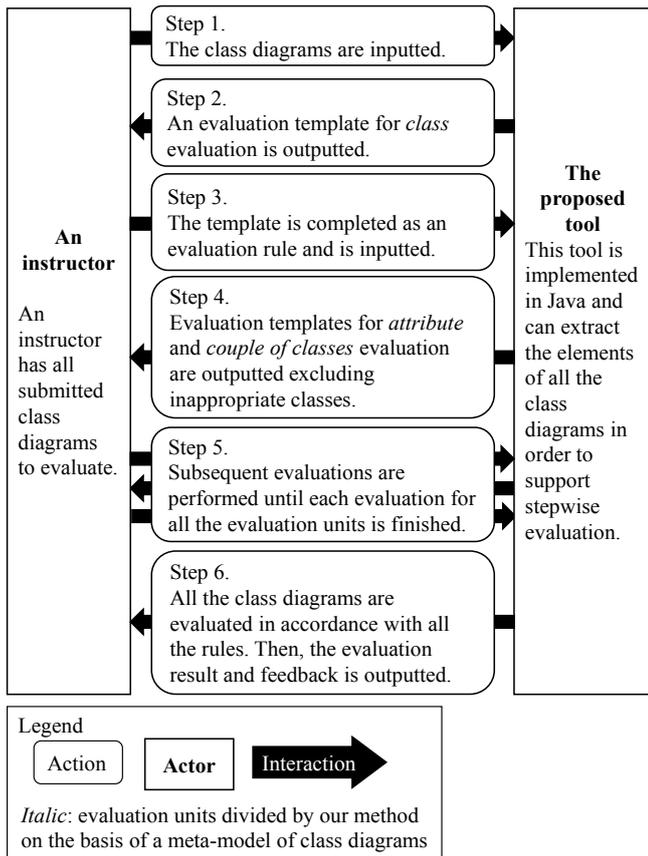


Figure 1: Overview of the proposed method.

Class name	NEAN	NECN	...	No Typo	No Other
ProductInventorySheet					
Product					
Camdy					
RiceBall					
InventoryManagementSystem					

Figure 2: Sample of the evaluation template.

[5], a UML modeling tool, and submitted by a learner. The files are then collected to a directory by the instructor, and the directory path as the collection of class diagrams is inputted to the tool. No file contains diagrams unrelated to the assignment that was given to the learners.

Step 2. Fig. 2 shows a sample of the evaluation template generated by the tool. This template can be automatically generated from all the learners’ class diagrams by aggregating the class diagrams’ information. An evaluation template is defined as a template of evaluation rules. Hence, the table representing the template in Fig. 2 does not include any evaluation marks.

The subject of this sample is an inventory management system for a small grocery store. According to this sample, one of the learners created the classes “Product,” “Camdy,” “RiceBall,” and so on. However, not all the classes in the table were always written by one learner, because this table was generated by aggregat-

ing class diagrams. It should be noted that “Camdy” is a typographical error, written instead of “Candy.” In fact, some learners make typographical errors in the diagram’s elements.

We call a row of this table an evaluation rule fragment. In an evaluation rule fragment, the part showing the class diagram information is called an evaluation target. The evaluation target consists of a single element or a combination of the elements in the class diagrams. The remaining part of an evaluation rule fragment shows the evaluation items. Evaluation items are specifically determined by a target evaluation unit. Fig. 3 shows all the evaluation units (e.g., “Class” and “Attribute”) and the dependency between the units.

The table in Fig. 2 was generated for the evaluation unit “Class.” Thus, this table has the evaluation items “NEAN” (no excessive abstract name), “NECN” (no excessive concrete name), “No Typo” (no typographical errors), “No Other” (no other errors), and so on. The dependency is explained in Step 4 in detail.

Multiple learners frequently submit similar answers to an assignment. Hence, in our opinion the efficiency of the evaluation is enhanced by deduplication of the evaluation targets. Fig. 4 shows the basic idea of extracting evaluation targets from class diagrams.

In Fig. 4, whereas the class diagram on the left hand side contains two classes, that on the right hand side contains four classes. First, our tool extracts single elements (e.g., a class name) or combinations of elements (e.g., an attribute name and the contextual class name) for each class diagram. The extracted single elements or combinations of elements are the candidate evaluation targets. Then, our tool compares each of the candidates in the different class diagrams. If the candidate is duplicated in different class diagrams, our tool deduplicates it in the process of generating the evaluation templates.

Furthermore, the dependency is used for reasonably reducing the candidate evaluation targets in this process. The details of using the dependency are presented in Step 4.

Step 3. The instructor completes an evaluation rule by entering the mark poor (N) in the evaluation template. In other words, an evaluation template after the mark has been entered is called an evaluation rule. Specifically, the instructor enters the mark (N) in the cell of an evaluation item in an evaluation rule fragment if the evaluation target does not satisfy the evaluation item.

Fig. 5 shows a sample of evaluation rules that corresponds to Fig. 2. The blank cells in the rule represent that the evaluation targets satisfy the corresponding evaluation items from the instructor’s viewpoint. Thus, the instructor can easily evaluate the evaluation targets.

Step 4. The tool automatically generates a next template on the basis of the result of the previous evaluation. Fig. 6

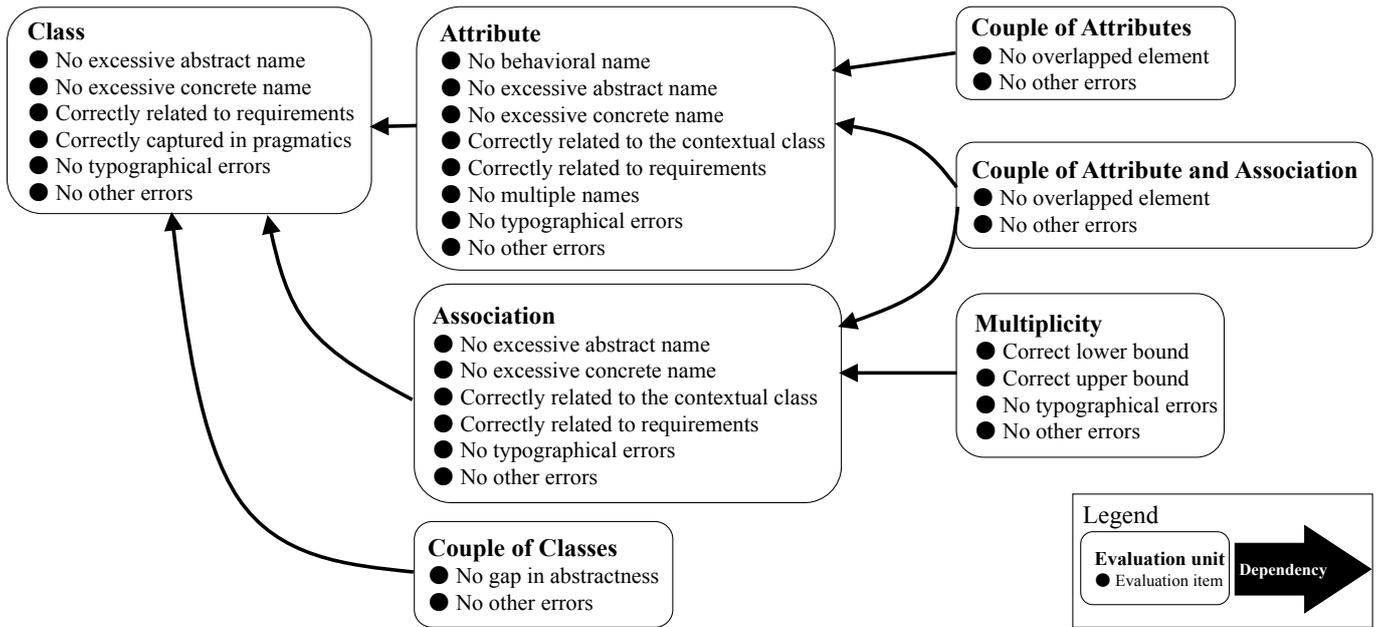


Figure 3: Evaluation units and the dependency between the units.

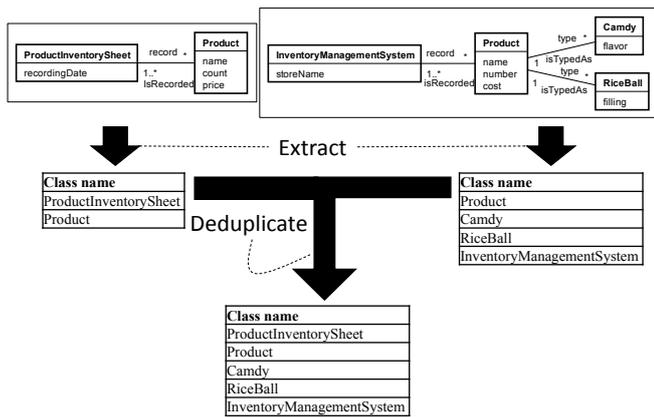


Figure 4: Basic idea for generating evaluation templates.

Class name	NEAN	NECN	...	No Typo	No Other
ProductInventorySheet					
Product					
Camdy		N		N	
RiceBall		N			
InventoryManagementSystem	N				

Figure 5: Sample of the evaluation rule.

shows a technique for stepwise evaluation of class diagrams. According to the advanced idea of our method, the elements (e.g., attributes) having a contextual element (e.g., a class) that was evaluated as poor, are not evaluated, since they cannot be precisely evaluated. Hence, such elements are excluded from the evaluation targets in order to reasonably reduce the instructor’s time cost.

At the top of Fig. 6, a sample evaluation template for class evaluation is shown. In the middle of the figure,

an evaluation template is shown where the instructor has entered one or more marks in the three evaluation fragments that are shaded gray.

Then, the tool generates a next evaluation template on the basis of the previous evaluation rule, shown at the bottom of the figure. In this case, an evaluation template for attributes contains the evaluation targets (e.g., attributes), the contextual elements (e.g., classes) of which were evaluated as not poor. In other words, the classes “Camdy,” “RiceBall,” and “InventoryManagementSystem” are excluded from the evaluation template for attributes. Such exclusions are performed based on our advanced idea mentioned above.

The order of generating the evaluation templates is defined as dependency in Fig. 3. For instance, the evaluation template for attributes is generated after creating the evaluation rules for classes. In Fig. 3, the target of an arrow represents a dependee and the source of the arrow represents a depender. Each dependee and depender constitutes an evaluation unit. Hence, the dependee units are evaluated prior to the depender units.

The evaluation items in every evaluation unit are redefined, extended, or excerpted based on the results of our previous study [9]. Clearly, the evaluation items (but not the evaluation units) used in our method can be easily customized, because the meaning of every evaluation item is not interpreted by the tool.

Step 5. The instructor continues subsequent evaluations until all the evaluations in Fig. 3 are completed.

Step 6. When the instructor has completed all the evaluations, the tool outputs the resulting final results and feedback shown in Fig. 7.

Class Evaluation

Class name	NEAN	NECN	...	No Typo	No Other
ProductInventorySheet					
Product					
Camdy					
RiceBall					
InventoryManagementSystem					

↓ evaluate

Class name	NEAN	NECN	...	No Typo	No Other
ProductInventorySheet					
Product					
Camdy		N		N	
RiceBall		N			
InventoryManagementSystem	N				

↓ generate a next template
Attribute Evaluation

Class name	Attribute name
ProductInventory	recordingDate				
Product	name				
Product	count				
Product	price				
Product	number				
Product	cost				

Figure 6: Process of stepwise generating evaluation templates.

The results table at the top of Fig. 7 summarizes the evaluation of all the units. A learner can be identified as a row in this table. In addition, the table enumerates all the evaluation items in all the units.

One of four marks, including a blank, is entered in the cell of each evaluation item in this table. The first mark constitutes a blank cell; that is, no symbol is entered. The blank cell means that all of the corresponding evaluation targets satisfied the evaluation item. In other words, none of the learner’s evaluation targets was marked as poor (N) by the instructor. The second mark is “X,” which means that one or more evaluation targets were marked as poor (N).

The third mark is “E,” which means that the diagram includes evaluation targets, but each of them contains

Learner	Class					Attribute				Association				Final result
	NEAN	NECN	...	No Typo	No Other	
A		X		X			E	E		-	-			X
B								X		X				X
C														
...

Feedback to A
 * The class “RiceBall” has an excessive concrete name.
 You should write the class name without any excessive concrete name.
 ...

Legend
 : All elements in the unit are correct.
 X : At least one element is incorrect.
 E : All the contextual elements are incorrect.
 - : No element to evaluate exists.

Figure 7: Final results and feedback automatically generated by the tool.

one or more contextual elements that were marked as poor (N). Such evaluation targets are not evaluated until all of their contextual elements have been evaluated as correct. The final mark is “-,” which means that the diagram has no evaluation target that should be evaluated regarding the evaluation item. While a final result becomes poor (N) if one or more evaluation targets was marked as poor (N), the final result becomes a blank cell if all the marks entered were either a blank or “-.”

Feedback messages can be automatically generated from the evaluation targets that were marked as poor (N). The learners should modify their diagrams until such messages are no longer outputted.

4 Experiments

4.1 Overview

Two experiments to answer two research questions (RQs) were conducted. As explained in Section 1, it seems that no existing tool is similar to our tool. Hence, the method presented in this paper (hereafter referred to as the extended method) was compared with our previous method [11] (hereafter referred to as the previous method) and a traditional one (hereafter referred to as the base method). The difference in the features of the extended and the previous method is that only the extended method employs the evaluation that utilizes the dependency shown in Fig. 3. In the traditional method, an instructor uses only the editor provided by Astah to confirm learners’ diagrams.

RQ 1. Does the extended method enhance the efficiency of evaluating class diagrams as compared to the previous and the base method?

RQ 2. Does the extended method adversely affect the final results that were obtained using the previous method?

In order to answer RQ 1, the evaluation process and product were quantitatively measured in an experiment (Exp. 1). In the case of the extended and previous methods, the time taken by the instructor to create all the evaluation rules was measured. In addition, the number of the evaluation rule fragments per evaluation unit was measured. In the case of the base method, the time taken by the instructor to manually evaluate all the class diagrams was measured. It was not possible to trace completely the manner in which the instructor evaluated which elements.

These measurements were already taken for the previous and the base method in our previous study. In Exp. 1, we newly conducted the evaluation for the extended method, and then, the results of applying the three methods were compared.

The preconditions of applying the three methods are explained as follows. Ninety-five class diagrams created by freshmen students at Shinshu University were used. Each diagram was written in natural language to satisfy certain requirements. The solution examples contain two or three classes, two or three attributes per class, and one or two associations in the diagram. In other words, the size of the examples is

Table 1: Number of evaluation rule fragments in the evaluation rules

Evaluation unit	Base method	Previous method	Extended method
Class	N/A	0	9
Attribute	N/A	90	86
Association	N/A	43	37
Multiplicity	N/A	61	43
Couple of Classes	N/A	13	10
Couple of Attribute and Association	N/A	175	118
Couple of Attribute	N/A	54	40
Total	N/A	436	343

Table 2: Time taken to evaluate the class diagrams

Base method	Previous method	Extended method
1 h 3 m and 19 s	27 m and 58 s	24 m and 32 s

very small. However, the submitted diagrams differed because frequently different students assigned different names to the elements.

Fig. 3 shows the evaluation items as improved after Exp. 1; a previous version of the evaluation items was used in Exp. 1. Specifically, the items “Correctly captured in pragmatics” and “No other errors” in the unit “Class,” “No excessive abstract name,” “Correctly related to the contextual class,” “No multiple names,” and “No other errors” in the unit “Attribute,” “Correctly related to the contextual class,” and “No other errors” in the unit “Association,” and “No other errors” in the rest of the units were not addressed. In addition, the items “Correct lower bound” and “Correct upper bound” in the unit “Multiplicity” were integrated into “Adequate multiplicity” in the previous version.

In the evaluation rules in Exp. 1, the mark of good (Y) was given to the cell of an evaluation item if the evaluation target satisfied the evaluation item. A smaller number of evaluation items was handled by the previous and base method than by the extended method: the number of the evaluation items in the previous and base method was 8 and in the extended one 17.

Although the instructor and all the class diagrams in Exp. 1 were the same as in the experiment in our previous study, the instructor had forgotten the previous evaluation results because a year had passed during which he had not used the tool. The instructor did not refer to any documents created in the experiment conducted in our previous study. The number of evaluation rule fragments per evaluation unit in the rule is shown in the row “Previous method” in Table 1.

To answer RQ 2, we used the extended method, reusing the evaluation rules that were obtained in our previous study. If the extended method did not change the final results that were obtained by the previous method, we assumed that the extended method does not adversely affect the results of the previous method. Hence, we applied the evaluation rules that were obtained by the previous method to the extended method. This experimental application is referred to as Exp. 2.

When the evaluation rules in the previous version were transformed into the rules in the new version, evaluation information was missing, since the number of evaluation items

in the previous and the new version differed. Hence, all the missing marks were completed with the mark good (Y), because these were not evaluated in the previous work. In Exp. 2, the target diagrams were the same as in Exp. 1.

4.2 Results

Each number of the evaluation rule fragments in Exp. 1 is shown in Table 1. The numbers of fragments for the base method were not available, since the evaluation was conducted without using the tool. The times taken to perform the evaluation are shown in Table 2. The “Total” time taken when the extended method was used was less than when the base method was used, although a greater number of evaluation items was included in the extended than in the base method. The final results of Exp. 2 showed no difference between the two methods.

4.3 Discussion

The results of Exp. 1 show that the answer to RQ 1 is “yes,” because the evaluation time when using the extended method is less than that when using the base and previous methods, although the number of the evaluation items in the extended method is more than in the previous method. Therefore, according to the results of Exp. 1, using the extended method instructors can more efficiently evaluate class diagrams than when using either the base or the previous method.

A problem related to the extended method is that the dependency between units is not adequately captured. For example, a dependency existed between “Association” and “Couple of Classes.” This dependency is bidirectional, which means that both units are always evaluated as poor when either one is evaluated as poor. We plan to extend our method to handle such bidirectional dependencies.

The results of Exp. 2 show that the answer to RQ 2 is “yes,” because the extended method with the complementary previous rules did not change the final results that were obtained in the previous study. Therefore, according to the results of Exp. 2, the extended method did not adversely affect the results of the previous method.

Consequently, the extended method has an advantage over the base method from the viewpoint of the efficiency of the evaluation. In addition, it is also completely effective as compared to the previous method.

5 Conclusion

This paper presented a rule-based method of stepwise evaluation of class diagrams. The method extends the previous method by employing the dependency between the evaluation units to stepwise evaluate class diagrams. Through two experiments, it was shown that the extended method can be expected to efficiently evaluate class diagrams without changing the evaluation quality in comparison with the previous method. As future work, we plan to improve the method for handling the bidirectional dependency and to extend our tool such that the evaluation items are customized, allowing reuse of existing rubrics or check lists (for example, [2], [15]).

Acknowledgment

This work was supported by JSPS KAKENHI Grant Numbers JP16H03074 and JP16K00094.

REFERENCES

- [1] Ali, N.H., Shukur, Z., Idris, S.: A design of an assessment system for uml class diagram. In: Proc. of the International Conference on Computational Science and its Applications 2007. pp. 539–546 (Aug 2007)
- [2] Ali, N.H., Shukur, Z., Idris, S.: Assessment system for uml class diagram using notations extraction. IJCSNS International Journal of Computer Science and Network Security 7(8), 181–187 (2007)
- [3] Andriyevska, O., Dragan, N., Simoes, B., Maletic, J.I.: Evaluating uml class diagram layout based on architectural importance. In: Proc. of the 3rd IEEE International Workshop on Visualizing Software for Understanding and Analysis. pp. 1–6 (2005)
- [4] Brandsteidl, M., Seidl, M., Kappel, G.: Teaching models @ big: On efficiently assessing modeling concepts. In: Educators' Symposium @ Models 2009 (2009)
- [5] ChangeVision: Astah, <http://astah.net/>
- [6] Demuth, B.: How should teaching modeling and programming intertwine? In: Proc. of the 8th Edition of the Educators' Symposium. pp. 32–35 (2012)
- [7] Gogolla, M., Vallecillo, A.: On explaining modeling principles with modeling examples: A classification catalog. In: Proc. of the 8th Edition of the Educators' Symposium. pp. 28–31 (2012)
- [8] Hasker, R.W.: Umlgrader: An automated class diagram grader. J. Comput. Sci. Coll. 27(1), 47–54 (2011)
- [9] Kayama, M., Ogata, S., Masamoto, K., Hashimoto, M., Otani, M.: A practical conceptual modeling teaching method based on quantitative error analyses for novices learning to create error-free simple class diagrams. In: Proc. of the 3rd International Conference on Advanced Applied Informatics. pp. 616–622 (2014)
- [10] Laforcade, P.: Towards a uml-based educational modeling language. In: Proc. of the Fifth IEEE International Conference on Advanced Learning Technologies. pp. 855–859 (July 2005)
- [11] Miyajima, K., Ogata, S., Kayama, M., Okano, K.: Tutoring support tool for rule-based scoring of class diagrams in uml modeling education (in japanese). In: Proc. of the 22nd Foundation of Software Engineering. pp. 169–174. Japan Society for Software Science and Technology (2007)
- [12] Moisan, S., Rigault, J.P.: Teaching object-oriented modeling and uml to various audiences. In: International Conference on Model Driven Engineering Languages and Systems. pp. 40–54. LNCS 6002, Springer (2009)
- [13] Ramollari, E., Dranidis, D.: Studentuml: An educational tool supporting object-oriented analysis and design. In: Proc. of the 11th Panhellenic Conference on Informatics. pp. 363–373 (2007)
- [14] Sun, Y., Yun, R.: Uml modeling for software system of edu-game. In: International Conference on Technologies for E-Learning and Digital Entertainment. pp. 395–404. Springer (2010)
- [15] Unhelkar, B.: Verification and validation for quality of UML 2.0 models. John Wiley & Sons (2005)

Session 4:
Multimedia Systems
(Chair: Katsuhiko Kaji)

TinyASIO: A C++ library for using ASIO with short source code.

Eiichi Takebuchi[†], Tomoki Kajinami^{††}, Ichiro Tokuhiro^{†††}, Haruo Hayami^{†††}

[†]Doctoral Course of Informatics, Kanagawa Institute of Technology, Japan

^{††}Okayama University of Science, Japan

^{†††}Kanagawa Institute of Technology, Japan

Abstract -

The audio interface is an acoustic equipment for converting an analog signal to a digital signal. By using ASIO advocated by Steinberg Company, the audio interface enables application and digital signal communication. When developing ASIO on Windows, it is necessary to call the COM interface with the Windows API, and development is not easy. As previous studies, ASIO SDK, RtAudio, PortAudio has been reported, but the development of ASIO is still not easy. In this paper, we propose a library to facilitate the development of ASIO. TinyASIO design hides Windows API, COM interface, ASIO specification. TinyASIO can develop ASIO without such deep knowledge and understanding. In the evaluation experiment, we evaluated two types of implementation examples in TinyASIO and previous studies. In TinyASIO, the Halstead volume was 2.6 times, the understanding/implementation time was 2.5 times, the maintainability index was 1.1 times better than the previous studies. In the library, by hiding elements necessary for ASIO implementation, audio applications can be implemented with fewer source code lines than previous studies. TinyASIO not only develops applications using the Audio Interface, but also improves the development of experimental systems.

Keywords: Audio Stream Input/Output, Audio Interface, C++ Library, Halstead Volume, Time to Implement

1 Introduction

In this paper, we propose a library (TinyASIO) that can use Audio Stream Input/Output (ASIO) with short program steps.

Instrument musical performers record their own performance with the Desktop Music software. Instrument musical performers play while listening to the performance. If the performance is delayed at this time, the performer becomes difficult to play.

The kernel mixer manages audio input/output on Windows. The kernel mixer is a system that manages audio input/output of multiple applications as one audio input/output. Since the kernel mixer simultaneously manages a plurality of voices, there is a problem that a delay is likely to occur.

ASIO is a common standard that enables audio input/output between an audio device and an application. Since ASIO can input/output audio between an audio

device and an application, it has the advantage of less delay than the kernel mixer.

ASIO development kit (ASIO SDK) is provided by Steinberg. ASIO SDK was very troublesome in the conventional library. TinyASIO encapsulates ASIO and improves maintainability and readability of source code. TinyASIO's contribution is the development efficiency of audio applications has improved.

Existing ASIO libraries have RtAudio[1] and PortAudio[2][3]. These libraries are cross-platform and support transmission standards other than ASIO. TinyASIO focuses only on the function of ASIO. TinyASIO is designed from the viewpoint of making ASIO easier to use.

In this paper, we describe a method to encapsulate ASIO to improve readability and maintainability. And, we evaluated the ease of development between TinyASIO and previous studies.

2 Current status of audio interface and ASIO

ASIO is a common standard of the audio interface provided by Steinberg[4][5]. Audio interface users use ASIO to transmits and receives digital signals between the application and the audio interface.

Steinberg distributes the ASIO SDK¹ as an ASIO development kit. Developers include the ASIO SDK to develop audio applications that utilize the audio interface. However, the ASIO SDK documentations do not exist. Developers need to understand from the ASIO SDK source code. If developers can read the document, developers can quickly understand the library. Developers take time to understand libraries that do not have documents.

Existing ASIO libraries include RtAudio[1] and PortAudio[2][3]. The problem with these libraries is API design with low readability and maintainability. These libraries are being developed with the aim of cross platform and cross libraries. These libraries have a complicated overall picture and execute all the constituent elements consistently.

In software development in recent years, developers need libraries that improve the readability and maintainability of source code[6][7]. It eliminates the complexity of ASIO and requires an environment where ASIO developers can easily code.

¹Steinberg Media Technologies: "ASIO SDK", <http://www.steinberg.net/en/company/developers.html>

3 TinyASIO

3.1 Library design

Knowledge of Windows API and COM is necessary for the development of ASIO. The development of ASIO will become easier by encapsulating advanced knowledge of libraries.

The relationship between TinyASIO’s class and actor is clear in the use case (Figure 1).

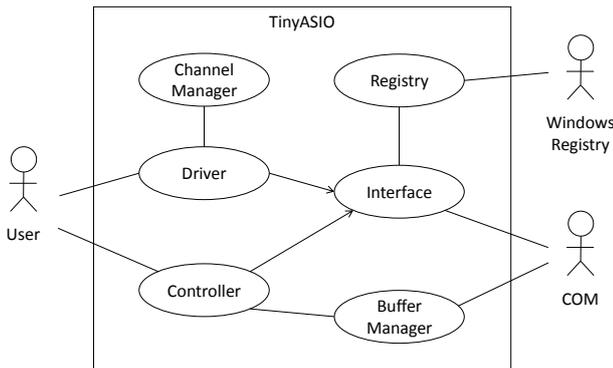


Figure 1: The usecase diagram of TinyASIO.

The use case shows that the class related to the registry and COM is encapsulated from the TinyASIO developer. TinyASIO encapsulates COM and Windows API from TinyASIO users in class design (Figure 2).

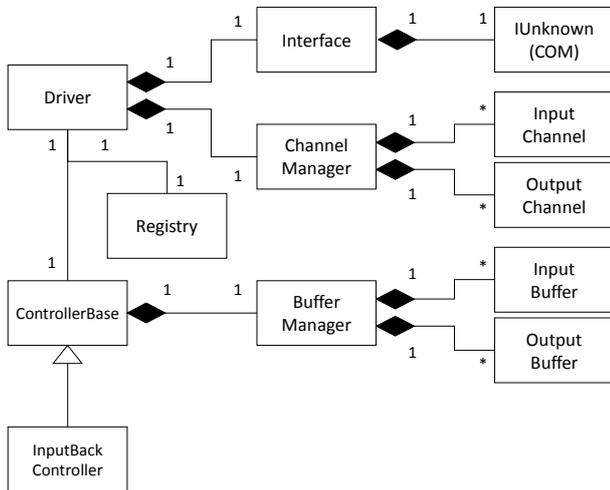


Figure 2: The class diagram of TinyASIO.

The Driver class encapsulates the Windows API and COM. The ControllerBase class controls ASIO. TinyASIO users can inherit the ControllerBase class and use ASIO. TinyASIO users can concentrate on developing applications using ASIO without using advanced knowledge.

3.2 Encapsulation of the registry and COM

Knowledge of the Windows API is required for operation of the registry and COM. TinyASIO encapsulates the registry and COM operations. Also, Windows appears conflicts when generating multiple COMs. TinyASIO avoids COM conflicts with a singleton pattern. The singleton pattern restricts the generation of class instances to one.

3.3 Synchronous communication between a computer and the audio interface

ASIO communicates between a computer and the audio interface. ASIO stores acoustic data in a temporary memory space called a buffer. Each channel of the audio interface has two buffers. ASIO switches buffers to save when one buffer is full (Figure 3).

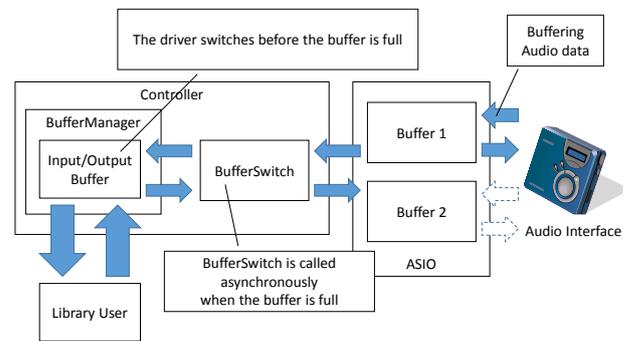


Figure 3: Relationship between BufferSwitch and Controller.

In the previous research, it is necessary to describe synchronously accessing temporary to memory space. In multithreading, asynchronous access to synchronous memory space has the risk of the deadlock. TinyASIO manages multithreading access of buffers with exclusive control.

3.4 Extension of TinyASIO

ASIO manages exchanges of sound data with the BufferSwitch function. The BufferSwitch function is callbacked when one of the two buffers becomes full. TinyASIO inherits the ControllerBase class and can define arbitrary controllers. Developers can easily extend the buffering process by defining the BufferSwitch function.

Developers can concentrate on developing buffering processes by using TinyASIO. In addition, developers don't need to manage sound data to input/output in BufferSwitch. By omitting the management of sound data, the responsibility of the source code is clarified.

4 The evaluation of the source code

4.1 The evaluation objective

In the evaluation of the source code, while evaluating the minimum implementation of ASIO using the library, evaluate the ease of development of the library. Ease of development here refers to source code that does not have much source code quantity, short understanding and implementation time, easy to maintain.

Evaluation targets are TinyASIO, ASIO SDK, RtAudio and PortAudio.

4.2 The evaluation index

We evaluated the quality of the source code[8] in the evaluation of the source code. The evaluation index of the source code is the number of program steps, Halstead Volume[9][10], the time required to program[11] and maintainability[12][13].

The program step is an executable program unit excluding comments and blank lines. Halstead Volume is a measure of the volume of a program. Halstead Volume increases with the vocabulary and number of words in the program. The time required program refers to the time it takes for the programmer to understand or implement the source code. Maintainability represents ease of maintenance.

4.3 Procedure for the evaluation of the source code

We considered two implementations for each library and evaluated each source code.

Implementation A evaluates whether TinyASIO is easier to implement than other libraries. Implementation A stores the data of the input buffer in another area and then writes the data of the input buffer to the output buffer.

Implementation B evaluates whether TinyASIO can implement the target source code just like any other library. Implementation example B randomly changes the data of the input buffer, stores the randomized data in another area, and writes the randomized data to the output buffer.

However, the implementation example of the ASIO SDK has conditions. The ASIO SDK needs to initialize COM with the Windows API. The initialization of COM is required CLSID. CLSID is a unique ID for COM. CLSID needs to refer to the registry of the ASIO driver. After looking up the corresponding registry from the ASIO driver name, refer to the registry where the CLSID is registered. After initializing the Windows API, initialize COM with CLSID. This sequence of processes is redundant. In the evaluation of the source code, Windows API and COM initializations are omitted.

4.4 The experiment environment

The operating system of the experiment environment is Windows 7 64 bits. Each implementation was com-

plied with VC++ 12.0. The development environment is Visual Studio 2013.

The ASIO SDK version is 2.3, RtAudio version is 4.1.1 and PortAudio version is V19.

The number of program steps and the number of lines were measured with Source Monitor V3.5².

In addition, we developed a system for evaluating the time required to program with Ruby 1.9.3.

4.5 The evaluation results

The evaluation results of implementation A and implementation B are shown below (Table 1).

The first line of the table is the result of implementation A, and the second line is the result of implementation B. The magnification represents the ratio of the previous study averages and TinyASIO. The numerical value in parenthesis is the ratio of the previous study and TinyASIO.

The evaluation result shows that TinyASIO is excellent with all indicators. Therefore, we considered being effective that the design of encapsulated ASIO and COM. TinyASIO encapsulates the initialization process of ASIO and COM and omits it.

4.6 Consideration

In the previous studies, initialization processing is indispensable. Developers are struggling writing source code in the initialization process. With this process, TinyASIO improves development efficiency in the early stages of development.

In TinyASIO, implementation B shows greater numerical values growth than implementation A. Compared with previous studies, TinyASIO has a big difference between implementation A and implementation B. This is because the ControllerBase class was derived and implemented. By simplifying the implementation of the inherited classes, we believe that TinyASIO will be a more easily developable library. Developers can be implemented in even a short period of time 2.5 times greater than previous studies with TinyASIO. TinyASIO has improved development efficiency and maintainability by encapsulating the initialization process of ASIO and COM.

Focusing on previous studies, the ASIO SDK is the smallest the time required to program in previous studies. ASIO SDK is a COM wrapper library, but it is suitable for minimal implementation. However the ASIO SDK requires Windows API and COM initialization. If developers do not have knowledge of Windows API and COM, we think it is difficult to use the ASIO SDK.

5 Conclusion

In this paper, we proposed TinyASIO to facilitate the development of ASIO. TinyASIO is a library that encapsulates ASIO and COM. TinyASIO improves the development efficiency of audio applications. TinyASIO was

²Jim Wanner: "Source Monitor V3.5", <http://www.campwoodsw.com/sourcemonitor.html>

Table 1: Evaluation results of implementation examples.

	ASIO SDK	RtAudio	PortAudio	TinyASIO	Average	Magnitude
Program Steps	63 (7.9)	58 (7.3)	73 (9.1)	8	64.7	8.1
	63 (2.6)	57 (2.4)	71 (3.1)	24	63.7	2.7
Halstead Volume	2,877 (9.7)	2,951 (9.9)	3,206 (9.1)	297	3011.3	10.1
	2,892 (2.6)	2,871 (2.6)	3,164 (3.0)	1,124	2975.7	2.6
Time Required to Program	1,917 (13.7)	4,918 (35.1)	3,116 (22.3)	140	3317.0	23.7
	2,169 (2.5)	4,784 (5.5)	3,076 (3.5)	874	3343.0	3.8
Maintainability	59.2 (1.7)	11.0 (9.1)	66.9 (1.5)	100.0	45.7	2.2
	59.1 (1.3)	14.8 (5.1)	68.9 (1.1)	75.6	47.6	1.6

2.6 times smaller in the Halstead volume than in previous studies.

TinyASIO's task is to further shorten the program steps. If we can describe the necessary requirements for audio applications short, we can expect to further improve the development efficiency of audio applications.

We are expecting more audio applications to be developed. TinyASIO can develop audio applications more easily than using ASIO SDK. TinyASIO can be used for study as well as audio applications[14][15][16][17]. We are expecting TinyASIO to promote the development of audio applications.

Acknowledgment

TinyASIO is published under the GNU General Public License³. I deeply appreciate the many who cooperated in debugging.

REFERENCES

- [1] Scavone, G. P.: RtAudio: A Cross-Platform C++ Class for Realtime Audio Input/Output, *Proceedings of the International Computer Music Conference*, pp. 196–199 (2002).
- [2] Bencina, R.: PortAudio and media synchronisation-it's all in the timing, pp. 13–20 (2003).
- [3] Ross, B. and Burk, P.: PortAudio an open source cross platform audio API, pp. 263–266 (2001).
- [4] Takazawa, H. and Tokuhiko, I.: Consideration about the validity of ASIO(Audio Stream Input Output), *EA Section, IEICE*, Vol. 108, No. 115, pp. 31–36 (2008).
- [5] Takebuchi, E. and Hayami, H.: Development of library to realize real-time audio processing with 5 lines, *The Special Interest Group Technical Reports of IPSJ, GN*, Vol. 9, No. 40, pp. 1–8 (2015).
- [6] Boswell, D. and Foucher, T.: *The art of readable code*, O'Reilly Media, Inc. (2012).
- [7] Kato, T., Kato, A., Okamura, N., Kanai, T., Suzuki, R. and Shirai, Y.: Musasabi: 2D/3D intuitive and detailed visualization system for the forest, *Proceedings of the ACM SIGGRAPH 2015 Posters*, ACM, p. 79 (2015).
- [8] Jiang, Y., Cuki, B., Menzies, T. and Bartlow, N.: Comparing design and code metrics for software quality prediction, *Proceedings of the 4th international workshop, Predictor models in software engineering on*, ACM, pp. 11–18 (2008).
- [9] Welker, K. D.: The software maintainability index revisited, *CrossTalk*, Vol. 14, pp. 18–21 (2001).
- [10] Halstead, M. H.: *Elements of software science*, Vol. 7, Elsevier New York (1977).
- [11] Halstead, M. H.: Toward a theoretical basis for estimating programming effort, *Proceedings of the 1975 annual conference*, ACM, pp. 222–224 (1975).
- [12] Land, R.: Measurements of software maintainability, *Proceedings of the 4th ARTES Graduate Student Conference*, pp. 1–7 (2002).
- [13] Heitlager, I., Kuipers, T. and Visser, J.: A practical model for measuring maintainability, *Proceedings of the Quality of Information and Communications Technology (QUATIC)*, IEEE, pp. 30–39 (2007).
- [14] Takebuchi, E., Inoue, H., Tokuhiko, I. and Hayami, H.: Proposal of pitch estimation algorithm by octave compression of pitch information, *Multimedia, Distributed, Cooperative, and Mobile Symposium 2015*, Vol. 33, No. 12, pp. 1103–1109 (2015).
- [15] Takebuchi, E., Kajinami, T., Tokuhiko, I. and Hayami, H.: Proposal of fast chord recognition method using bass pitch, No. DS-12, pp. 1840–1847 (2015).
- [16] Takebuchi, E., Kajinami, T., Tokuhiko, I. and Hayami, H.: Using MIDI Sequence's Standard Deviation Classifying a Fundamental Frequency And Extracting a Picking Time and a Playing Time, *The Special Interest Group Technical Reports of IPSJ, MUS*, Vol. 113, No. 16, pp. 1–5 (2016).
- [17] Takebuchi, E., Kajinami, T., Tokuhiko, I. and Hayami, H.: Examination of the Bass Pitch Detection Algorithm in the Electric Bass, *International Workshop on Informatics (IWIN2016)*, pp. 197–202 (2016).

³<https://github.com/GRGSIBERIA/TinyASIO>

A Method for Video Advertisement Insertion with Audience Comments on Action Game Videos

Yoshia Saito*

*Faculty of Software and Information Science, Iwate Prefectural University, Japan
y-saito@iwate-pu.ac.jp

Abstract – We have proposed a video advertisement insertion method which does not interfere with video viewing utilizing audience comments in video sharing services. The video advertisement insertion method, however, could apply for the videos which had only a few shot boundaries. Meanwhile, action game videos are very popular contents in video sharing services nowadays. We conducted an experiment to verify the effectiveness of the previous method using action game videos in niconico. As the result, we found the previous method could not be applied for the most videos. It was necessary to improve the method in order to handle various action game videos. In this paper, we analyze details of the action game videos such as changes of the number of comments. Based on the analysis, we propose a new video advertisement insertion method which can handle various action game videos. We also evaluate whether the proposed method enables audience to view the video with comfort even if the video advertisement is inserted.

Keywords: Video Advertisement, Insertion Algorithm

1. INTRODUCTION

High-speed Internet access is available for a wide range of people and the people can watch high-quality video via the Internet. There are various video sharing services which enable people to upload and watch videos. The typical ones in Japan are YouTube [1] and niconico [2]. In these video sharing services, people can upload and watch videos for free but advertisements are displayed instead. While there are various types of advertisements, mid-roll video advertisements become popular in recent years. The mid-roll video advertisements insert a video advertisement in the middle of viewing a video content like TV commercials. Adobe reports the mid-roll video advertisements are engaging commercials which have high completion rate [3]. The mid-roll video advertisements will be used even further in the next several years.

The mid-roll video advertisements, however, create disadvantage for viewing audience. If a video advertisement is inserted at the wrong time, the audience feel uncomfortable about the advertisement and it will reduce effectiveness of the advertisement. There is a psychological process model of advertisements, AIDMA [4] which is a process that leads consumers to purchase some products.

The process goes along “Attention”, “Interest”, “Desire”, “Memory” and “Action”. At first, consumers watch advertisements and aware of a products (Attention). Then, they are interested in it (Interest) and desire to get it (Desire). They memorize the product (Memory) and purchase it at last (Action). If the audience feel uncomfortable about the video advertisement which is inserted at the wrong time, the process will stop at “Attention” phase and not proceed to “Interest” phase. Therefore, it is important to insert video advertisements at right timing.

We have proposed a video advertisement insertion method which does not interfere with video viewing utilizing audience comments in video sharing services [5]. The niconico, which is one of the video sharing service in Japan, enables audience to submit comments with playback time of the viewing video and displays the comments at the playback time as if they had a chat with the other audience. The comment can be regarded as metadata of the video scenes and utilized for analysis of the video content. Our method uses the comments to find a timing not to interfere with video viewing so that it can insert a video advertisement to the video at the right timing. However, it has a limitation which applies only to videos with a few shot boundaries.

Meanwhile, action game videos are very popular contents in video sharing services. They are videos of playing action games which are generated by users and generally include a lot of shot boundaries. It concerns the effectiveness of the previous method for action game videos. In this paper, we verify the effectiveness of the previous method using action game videos in the niconico and analyze characteristics of the action game videos such as changes of the number of comments. Based on the analysis, we propose a new video advertisement insertion method which can handle various action game videos.

The paper is organized as follows. In the next section, we describe related work of scene analysis using video metadata and interactive advertisement. In section 3, we describe the detail of our previous work. Then, we conduct an experiment to verify the effectiveness of the previous method in section 4. We analyze characteristics of the action game videos and propose a new video advertisement insertion method in section 5. In section 6, we evaluate the new method compared with the previous method and whether the proposed method enables audience to view the video with comfort even if the video advertisement is

inserted. Section 7 gives some conclusions and our future work.

2. RELATED WORK

Interactive television (iTV) [6-8] is a research area to give interactive features to television in order to break away from a traditional one-way broadcast service model. The iTV provides audience who watch the TV with various interactive functions with shared experiences. One of the iTV service is social TV [9, 10]. The social TV enables people to watch same TV content with the other audience such as friends and family via network and makes the viewer aware of context of other viewers. Typical social TV systems have chat functions such as text chat and voice chat to let audiences communicate with each other so that they have shared experiences. Internet broadcasting services such as YouTube and niconico have these features and become popular in recent years. There are many researches of scene analysis using metadata of the video [11-13], these researches typically use predefined metadata described by its producer. In our research, we use comments as metadata which is described by audience.

There are studies of interactive advertising to provide interactivity to the advertising [14-17]. The interactive advertising allows selecting appropriate ads according to the viewers and changing video length and display methods. Our research is regarded as one of technologies for interactive advertising. Tao Mei et al. [18] proposed a scheme of appropriate video ad insertion for online videos. In this research, the appropriate timing for video ad insertion is determined detecting an unattractive video cut. The unattractive video cut is detected by importance of the scene audio-visually. Our research uses audience comments to detect appropriate timing for video ad insertion.

3. PREVIOUS METHOD

We proposed a method which realized comfortable viewing of videos for audience in video sharing services even if video advertisements are inserted to the shared videos. We presumed the appropriate insertion timing for video advertisements had relation to number of audience comments per second because the comments represent how the audience was interested in the commented scene and the video advertisement should be inserted around less interesting scenes not to disturb their video viewing.

We studied an appropriate insertion timing for video advertisements by interviewing 20 students in our university. In the experiment, we chose 4 videos (video 1-4) which were shared in niconico and had 10,000 comments. The lengths of these videos are 5 or 10 minutes because a research reports 98% of videos in YouTube are less than 10 minutes [19]. The video 1 and 2 have 5 and 10 minutes respectively and over 30 shot boundaries. The video 3 and 4 have 5 and 10 minutes respectively and less than 10 shot boundaries. We asked 20 students in our university who had watched videos in video sharing services to answer best insertion timing for the chosen videos as its playback time.

From this experiment, we found timing on a shot boundary where context of the scenes is changed and the shot boundary

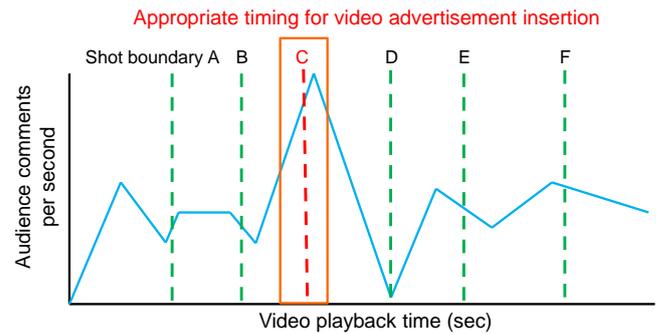


Figure 1: The previous method for video advertisement insertion using audience comments

had maximum variance of number of audience comments is appropriate to insert the video advertisement among all shot boundaries of the video if the shot boundaries were less than 10. Figure 1 shows how to find an appropriate timing in the previous method. Meanwhile, we did not see this characteristic in video 1 and 2 which had more than 30 shot boundaries.

4. EXPERIMENT

We conduct an experiment to verify the effectiveness of the previous method using action game videos in niconico. Game playing videos are popular contents in recent video sharing service and action game is most popular game genre in Japan. For this reason, we use action game videos in this experiment.

4.1. METHODOLOGY

At first, we select 6 action game videos (video 1-6) which are most viewed in niconico for the experiment and collect audience comments of these videos. We ask 20 subjects to answer 3 timings where they accept viewing video advertisement from 1st place to 3rd place. The precondition to select the timings is as follows.

1. The video advertisements are 15 seconds.
2. The video advertisements are your interested ones.
3. The video advertisements are played stopping the watched video like TV commercials.
4. The video advertisements are inserted to the video without start time and end time.

Then, we give score to the timings. Each timing of 1st place is 3 point, 2nd place is 2 point and 3rd place is 1 point. We calculate total of the points and derive appropriate timings from 1st place to 10th place in the video.

4.2. RESULT

We compare these timings with timings which are derived by the previous method. Table 1 shows the results. We

Table 1: Estimation accuracy of the previous method

	Timing derived by the previous method
Video 1	×
Video 2	×
Video 3	×
Video 4	×
Video 5	×
Video 6	corresponds with 1st place

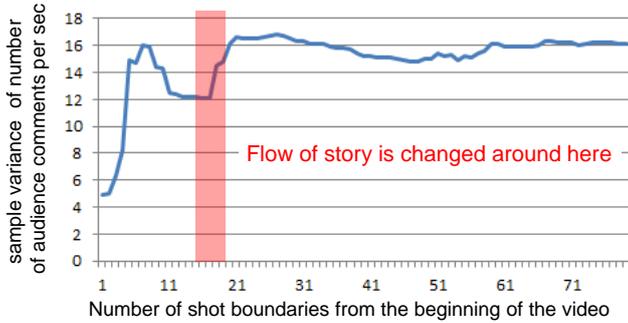


Figure 2: Relationship between change in sample variance of number of audience comments per second and timing when flow of the story is changed.

found our previous method can be applicable to only video 1. The action game video have a lot of shot boundaries and our previous method is not effective to these videos.

The problem of the previous method is not to consider the relationship among context of scenes through the whole of the video. The previous method considers context around each shot boundary independently and not enough to grasp flow of the story. Therefore, we should improve the previous method so that it can estimate when flow of the story is changed analyzing audience comments through the whole of the video.

5. PROPOSED METHOD

In this section, we analyze characteristics of the action game videos and propose a new video advertisement insertion method

5.1. ANALYSIS

To improve our previous method for video advertisement insertion, we analyze audience comments of the videos which were used in the experiment. We study whether a shot boundary where flow of the story is changed can be estimated by analyzing characteristics of audience comments. To find the shot boundary where flow of the story is changed, we adopt 2 steps approaches as follows.

1. Estimating several shot boundaries roughly which have a possibility include a shot boundary where flow of the story is changed.
2. Choosing a shot boundary which does not interfere with video viewing among the estimated shot boundaries.

Table 2: Variance of number of audience comments for each shot and the appropriate shot boundaries around flow of story is changed in video 1-6.

Video 1		
	Variance of number of audience comments for each shot	
Shot 1 (6:00-6:26)	8.81	
Shot 2 (6:26-6:28)	2.00	
Shot 3 (6:28-6:32)	6.80	
Shot 4 (6:32-7:05)	11.10	
Shot 5 (7:05-7:10)	5.67	← 4th place
Shot 6 (7:10-7:27)	43.54	← 3rd place
Shot 7 (7:27-8:19)	12.88	

Video 2		
	Variance of number of audience comments for each shot	
Shot 1 (6:25-6:50)	29.76	
Shot 2 (6:50-7:00)	32.96	← 2nd place
Shot 3 (7:00-7:02)	13.56	← 8th place
Shot 4 (7:02-11:00)	60.32	
Shot 5 (11:00-11:15)	3.47	← 5th place

Video 3		
	Variance of number of audience comments for each shot	
Shot 1 (5:58-6:10)	7.92	
Shot 2 (6:10-6:14)	2.16	
Shot 3 (6:14-6:26)	8.70	
Shot 4 (6:26-6:41)	1.73	
Shot 5 (6:41-6:56)	15.87	
Shot 6 (6:56-7:08)	27.41	← 2nd place
Shot 7 (7:08-7:34)	15.25	

Video 4		
	Variance of number of audience comments for each shot	
Shot 1 (7:54-7:57)	4.25	
Shot 2 (7:57-9:04)	47.42	
Shot 3 (9:04-9:19)	10.40	
Shot 4 (9:19-9:33)	21.40	← 7th place
Shot 5 (9:33-10:14)	67.72	← 1st place
Shot 6 (10:14-10:17)	7.50	← 10th place
Shot 7 (10:17-10:31)	19.00	← 10th place

Video 5		
	Variance of number of audience comments for each shot	
Shot 1 (1:44-1:58)	5.97	
Shot 2 (1:58-2:08)	3.88	
Shot 3 (2:08-2:18)	3.16	
Shot 4 (2:18-2:27)	6.25	
Shot 5 (2:27-2:56)	7.34	
Shot 6 (2:56-3:33)	4.07	← 1st place
Shot 7 (3:33-4:06)	89.43	

Video 6		
	Variance of number of audience comments for each shot	
Shot 1 (1:29-1:34)	7.81	
Shot 2 (1:34-1:47)	97.49	← 5th place
Shot 3 (1:47-2:21)	18.88	← 7th place
Shot 4 (2:21-2:26)	242.81	
Shot 5 (2:26-2:29)	130.69	
Shot 6 (2:29-2:31)	0.67	
Shot 7 (2:31-2:45)	141.81	← 9th place

At first, characteristics of audience comments is analyzed in order to estimate several shot boundaries roughly. From this analysis, we find a relationship between change in sample variance of number of audience comments per second and timing when flow of the story is changed. Figure 2 shows the change in sample variance of video 1. The x-

axis shows the change in sample variance of number of audience comments per second and the y-axis shows number of shot boundaries from the beginning of the video. There is a shot boundary where flow of story is changed around the red area in the figure. From this figure, we can see the flow of story is changed around time when the sample variance rises for the second time. This feature is also shown in the other videos.

Secondly, we need to find an appropriate shot boundary among the shot boundaries which are estimated by the change in sample variance. To analyze the characteristics of the audience comments, we focus on 7 shot boundaries around time when the sample variance rises for the second time. Table 2 shows the relationship between variance of number of audience comments for each shot and the appropriate shot boundaries. There are 7 shots centered at the time when the sample variance rises for the second time. Note that there are only 5 shots in video 2 because it is near the last shot when the sample variance rises for the second time. The first column of the table shows the shot number with the time period. The second column shows the variance of number of audience comments for each shot. The arrows mean the appropriate shot boundaries with rank order if they are present in the 7 shot boundaries. From the table, we can see all cases of 6 videos include the appropriate shot boundaries at least one shot. We can also find that maximum difference of the variance of number of audience comments between two adjacent shots coincides with one of the existed appropriate shot boundaries among the 7 shots.

5.2. NEW METHOD

Based on the two findings of the experiment, we propose a new method for video advertisement insertion which can be applicable to the action game videos even if they have a lot of shot boundaries. Note that the shot boundaries are assumed to be already-known using the existed methods. The new method is as follows.

1. Getting 10,000 audience comments of the video.
2. Detecting shot boundaries of the video.
3. Calculating sample variance of number of audience comments per second from first shot boundary to last shot boundary.
4. Omitting the first shot and the last shot from the following process.
5. Detecting a shot boundary which is first with lowest sample variance after the first boundary with maximum sample variance from start of the video.
6. Referring 7 shots centered at the detected shot boundary.
7. Calculating variance of number of audience comments per second for each shot.
8. Calculating difference of the variance between two adjacent shots.
9. Finding maximum difference of the variance.
10. Choosing the shot boundary for video advertisement insertion.

Table 3: Effectiveness of the previous method and the new method

	Previous method	New method
Video A	x	corresponds with 1st place
Video B	x	corresponds with 1st place
Video C	x	corresponds with 3rd place
Video D	corresponds with 2nd place	corresponds with 8th place
Video E	x	corresponds with 7th place
Video F	corresponds with 3rd place	x
Video G	x	corresponds with 4th place
Video H	x	corresponds with 5th place
Video I	x	corresponds with 2nd place
Video J	corresponds with 7rd place	corresponds with 1st place
Video K	x	corresponds with 1st place
Video L	x	corresponds with 7th place

This method requires 10,000 audience comments. It is provisional value and optimum value should be studied. In this paper, however, it is out of scope.

6. EVALUATION

We evaluate effectiveness of the new proposed method compared with previous method using other action game videos. We also conduct a subjective evaluation if the new method can choose an appropriate timing of video advertisement insertion for the audience and it does not interfere with their video viewing.

6.1. ESTIMATION ACCURACY

We evaluate the effectiveness of the new proposed method in the same manner as the previous experiment. We select 12 action game videos (video A-L) which are most viewed in niconico for the experiment and collect audience comments of these videos. We ask 20 subjects to answer 3 timings where they accept viewing video advertisement from 1st place to 3rd place. The precondition is also same as the previous experiment. Then, we gives score to the timings. Each timing of 1st place is 3 point, 2nd place is 2 point and 3rd place is 3 point. We calculate total of the points and derive appropriate timings from 1st place to 10th place in the video.

Table 3 shows the result. The previous method can be only applicable to 3 videos (video D, F, J). The new method can be approximately applicable to all videos without video F. This result shows the new proposed method can estimate an appropriate timing for video advertisement insertion effectively more than the previous method. Although the new method cannot estimate an appropriate timing of video F, the number of the shot boundaries in the video is less than 10. The previous method can be applicable if the number of the shot boundaries. Therefore, we need to use the methods properly according to the number of shot boundaries.

6.2. SUBJECTIVE EVALUATION

We evaluate whether audience can view the video without interruption by the video advertisement at the timing when the new method choose. We prepare 3 videos (video 1-3)

	Group 1 13 subjects	Group 2 14 subjects	Group 3 14 subjects	Group 4 13 subjects
Video 1	1st place timing New method	5th place timing	8th place timing	Timing of previous method
Video 2	1st place timing	5th place timing New method	8th place timing	Timing of previous method
Video 3	1st place timing	5th place timing	8th place timing New method	Timing of previous method

Figure 3: 4 groups for the subjective evaluation

which were used in the evaluation for estimation accuracy. Each video has appropriate timings for starting video advertisement from 1st place to 10th place which were calculated in section 6.1. We use the 1st, 5th and 8th place timing of the appropriate timing for the evaluation. The new method can estimate 1st place timing in the video 1, 5th place timing in the video 2 and 8th place timing in video 3. We obtain the cooperation of 54 students who have used video sharing services for research subjects in this evaluation. The subjects are divided into 4 groups as shown in Figure 3. There are 13 subjects in Group 1 and they watch the videos which is inserted a video advertisement of 15 seconds at 1st place timing. There are 14 subjects in Group 2 with video advertisement at 5th place timing. There are 14 subjects in Group 3 with video advertisement at 8th place timing. There are 13 subjects in Group 4 and they watch the videos which is inserted a video advertisement of 15 seconds at timing when the previous method estimates. After watching the videos, we ask the subject if the video advertisement interrupts video viewing on a scale of one (uncomfortable) to five (comfortable) by a questionnaire.

Figure 4-5 shows the results of video 1-3. From these figures, we can see largest number of subject answer it was comfortable when the video advertisement was inserted at the timing which is estimated by the new method. We found audience could watch the videos comfortably even if the timing estimated by the new method was low in rank.

Figure 7-9 shows the results comparing with the previous method in video 1-3. From these figures, the new method marks higher average score than that of the previous method in all videos. From figure 9, we found that the audience could watch the videos comfortably even if the timing estimated by the new method is lower in rank than the previous method.

From these result, our new method can estimate an appropriate timing for video advertisement insertion. Audience can watch videos comfortably without feeling of interruption of their video viewing. In addition, the new method is more effective than the previous method.

7. CONCLUSION

We proposed a new method for video advertisement insertion based on analyzing characteristics of audience comments. The new method can estimate an appropriate timing more accurately and enables audience to watch videos more comfortably without feeling of interruption of

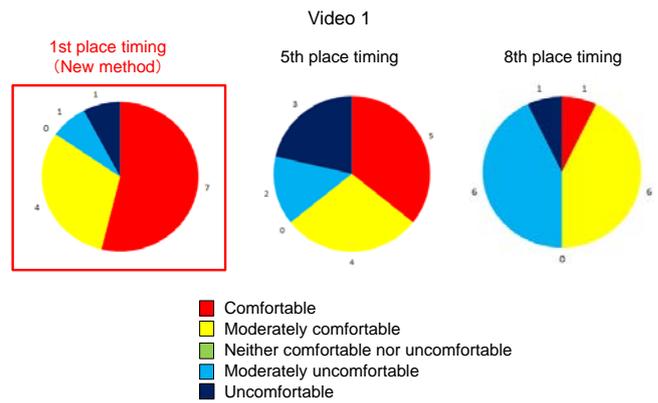


Figure 4: Evaluation result of video 1

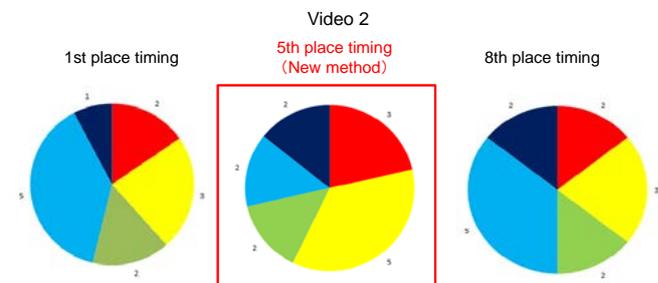


Figure 5: Evaluation result of video 2

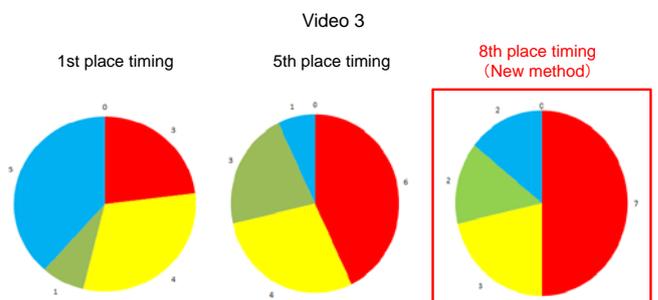


Figure 6: Evaluation result of video 3

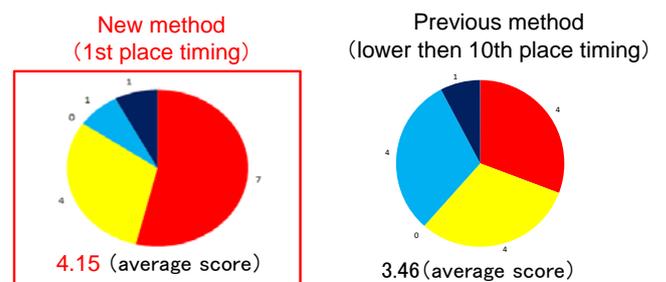


Figure 7: Comparing with previous method in video 1

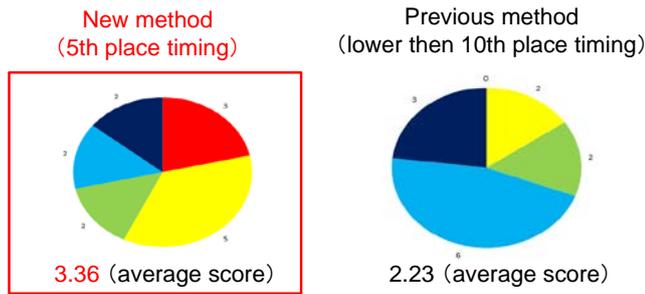


Figure 8: Comparing with previous method in video 2

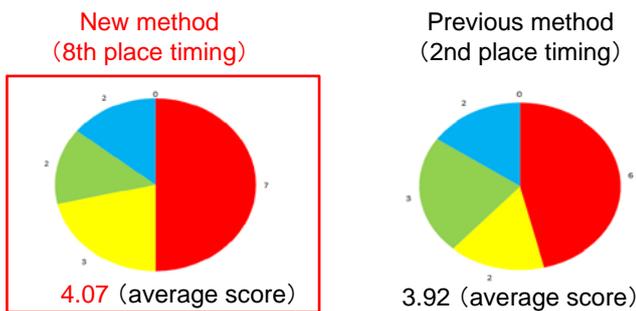


Figure 9: Comparing with previous method in video 3

their video viewing by the video advertisement insertion than the previous method. Note, however, that the previous method should be used if the number of the shot boundaries are less than 10.

In the evaluation, we found there were difference of the appropriate timing for video advertisement insertion with respect to each subject. The insertion timing should adapt to each subject differently. Moreover, our method needs a lot of audience comments and cannot be applicable to videos without comments. We will develop other method using sensor devices.

REFERENCES

[1] YouTube, <https://www.youtube.com/>
 [2] niconico, <http://www.nicovideo.jp/>
 [3] 2012 adobe digital video advertising report, https://blogs.adobe.com/primetime/files/2013/11/Monetization-Report_FINAL1.pdf
 [4] S. Roland Hall : Retail advertising and selling , The History of advertising: 40 major books in facsimile (1985).
 [5] Yoshia Saito and Yuko Murayama: Implementation of an Internet Broadcasting System with Video Advertisement Insertion Based on Audience Comments, International Conference on P2P, Parallel, Grid, Cloud and Internet Computing (2010).
 [6] J. Jensen: Interactive Television – A Brief Media History, In Proc. of EuroITV2008, pp. 1-10 (2008).

[7] M. Ursu et al.: Interactive TV narratives: Opportunities, progress, and challenges, ACM TOMCCAP, Vol. 4, Issue 4 (2008).
 [8] P. Cesar and K. Chorianopoulos: Interactivity and user participation in the television lifecycle: creating, sharing, and controlling content, In Proc. of UXTV 2008, pp.125-128 (2008).
 [9] P. Cesar, K. Chorianopoulos and J. Jensen: Social television and user interaction, Computers in Entertainment (CIE), Vol. 6 Issue 1 (2008).
 [10] L. Oehlberg, N. Ducheneaut, J. Thornton, R. Moore and E. Nickell: Social TV: Designing for Distributed, Sociable Television Viewing, International journal of human computer interaction, Vol. 24, No. 2, pp. 136-154 (2008).
 [11] Chitra L. Madhwacharyula, Marc Davis, Philippe Mulhem, Mohan S. Kankanhalli: Metadata handling: A video perspective, ACM Transactions on Multimedia Computing, Communications, and Applications (TOMCCAP), pp. 358-388 (2006).
 [12] A. Jaimes, T. Echigo, M. Teraguchi and F. Satoh: IEEE International Conference on Image Processing, (2002).
 [13] Y. Takahashi, N. Nitta and N. Babaguchi: Video Summarization for Large Sports Video Archives, IEEE International Conference on Multimedia and Expo, pp. 1170-1173 (2005).
 [14] K. Ridsen, M. Czerwinski, S. Worley, L. Hamilton, J. Kubiniec, H. Hoffman, N. Mickel and E. Loftus: Interactive advertising: patterns of use and effectiveness, SIGCHI, pp. 219-224 (1998).
 [15] Jong Woo Kim and Stephen Du: Design for an Interactive Television Advertising System, Proceedings of the 39th Annual Hawaii International Conference on System Sciences, Vol. 2 (2006).
 [16] J. Lloyd: I-Ads - a new approach, European Conference on Interactive Television (2003).
 [17] Panagiotis Giotis, George Lekakos: Effectiveness of Interactive Advertising Presentation Models, EuroITV '09, pp.157-160 (2009).
 [18] Tao Mei: VideoSense-Towards Effective Online Video Advertising, ACM Multimedia'07, pp.1075-1084 (2007).
 [19] Xu Cheng, Cameron Dale, Jiangchuan Liu: Statistics and Social Network of YouTube Videos, IEEE IWQoS 2008, pp. 229-238 (2008).

WakWak Tube: An Asynchronous YouTube Co-Viewing System

Yuichi Bannai* and Yusuke Wakatsuki**

*Department of Information Media, Kanagawa Institute of Technology, Japan

** JISC SOFT Co., Ltd., Japan

bannai@ic.kanagawa-it.ac.jp

yusuke.wakatsuki@gmail.com

Abstract - Previously conducted research on social TV has enabled remote audiences to chat via voice and text while watching TV. In this study, we develop an asynchronous YouTube co-viewing system called WakWak Tube, which captures the movement of the viewer's body using a Microsoft Kinect sensor as a YouTube video plays, and displays the viewer as an avatar in an audience area on the screen in real time. A new viewer can look at other avatars as well as his or her own avatar synchronously with the video playback. In this research, we propose a system that enables viewers to foster a sense of connection with other avatars by selecting avatars recorded in the past and reconfiguring the audience area to including their own avatars.

Keywords: Asynchronous Co-Viewing YouTube Avatar Kinect

1 BACKGROUND

There have been numerous studies addressing social television to support text- and audio-chat communications for remote audiences [1] [2]. A co-viewing system that couples user comments and actions with video playback time and reproduces past comments and actions in sync with the playback time can make the user feel connected to other audience members across time and space. These services take advantage of the merits of the Internet. However, they ignore nonverbal information, which often plays a more vital role than language in in-person interactions.

In contrast, at public movie viewings or live concerts, we can experience a sense of presence and unity that cannot be obtained through online communication, but are restrictive in that they require gathering in the same place at the same time. Some systems that share verbal and nonverbal information between viewers of a video-sharing service, such as public TV, have been proposed, but such systems for reconstructing the viewer space to include the user has not yet been developed.

The WakWak Tube system is designed for situations in which viewers move their bodies while watching YouTube videos, as they might while watching music concerts and dance performances. We consider the YouTube player to be a stage, and set up audience areas from which audience members can watch the stage. In this space, where all viewers appear as avatars, we consider how to enable users to feel a sense of presence and connectedness by configuring the space and avatars. We implement interaction functions between user avatars, such as pushing away other avatars to

make enough space for the movement of one's own avatar. The audience area can display nine avatars, including the viewer's own avatar, and the viewer can select and display desired avatars using properties such as time, gender, and activity level.

2 RELATED WORKS

CollaboraTV [3] is a system that supports both synchronous and asynchronous viewers, in which viewers are represented as icon avatars forming a virtual audience at the bottom of the screen. Viewers can share avatar actions, such as raising hands, giving a thumbs-up, and displaying text messages with speech balloons. Consequently, the viewer can watch TV and feel connected to other viewers.

Rabbit [4], which is a popular co-viewing application, allows group video, voice, and text chat while watching a movie.

Shirokura et al. [5] developed a system called ExciTube that can share information about viewers' excitement while watching a video. Viewer excitement is displayed using an avatar in a viewer area below the video window. The degree of excitement is measured by skin conductance response using an electrode attached to the palm. However, using skin conductance response as a physiological signal to directly measure emotion provides insufficiently rich shared information.

Yoshida et al. [6] developed a system to capture the motion of a viewer watching a movie in front of a large screen using a Microsoft Kinect, and to superimpose the motion onto the movie. This system aims to provide present viewers a sense of unity with past audiences by displaying skeleton information from past and present audience members as stick-like avatars. However, the authors concluded that viewers were unable to feel a sense of unity because it was difficult to distinguish one's own avatar from others, and there was no interaction between avatars.

Audience Silhouettes for TV [7], developed by Vatavu, is a system that enables the visual representation of viewers' movements displayed in real time over television content. The Microsoft Kinect depth sensor is deployed to capture audience member's movement. Using this motion data, the audience's motion and posture can be shown as silhouettes. This system has the merit of conveying nonverbal information by sharing viewer silhouettes, connecting viewers of the content. However, its disadvantages include a lack of discussion of space design for displaying silhouettes and interacting with them

In online video viewing systems such as YouTube, viewers can only have the mean of leaving their comments as a communication channel because they watch the video asynchronously. However, we think that it is possible to convey nonverbal information among asynchronous users along with the common time axis based on the YouTube video playback time. The contributions of this research are as follows: (1) we propose the audience space where the body movements of past viewers as well as current viewer are synchronously displayed as avatars along with the video playback time; (2) we provide means for reconstructing this space by selecting desired past avatars; (3) we also provide feasible interaction means with the past avatars.

3 PROTOTYPE SYSTEM OVERVIEW

3.1 System Overview

Figure1 shows the usage of WakWak Tube. The user stands up in front of the PC screen on which a Kinect is placed. The viewer’s skeletal data, representing joint locations, is captured by a Kinect v2 depth sensor. Then the user’s avatar animation is displayed on the screen while the user moves his or her body while watching a YouTube video.

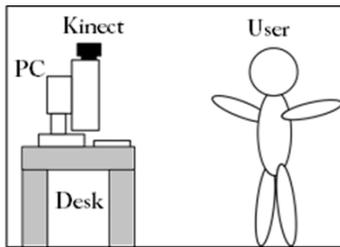


Figure1: Usage of WakWak Tube

3.2 System Usage

After starting up the system, the user inputs his or her login ID and gender on the panel shown in Figure 2 and enters WakWak Tube.



Figure 2: Login panel

Then, the search and selection panel, containing a text box for video search and two pull-down menus for avatar selection, appears (Figure 3). The avatar display area has two pull-down menus that determine avatars to be displayed along with the user’s own. In the “Display order” menu, “newest first” selects avatars by descending order of recorded time; “most active first” displays avatars by activity level, starting with highest.

When the user inputs a keyword, a list of video information acquired from YouTube is displayed in search results containing the video title and the beginning of its

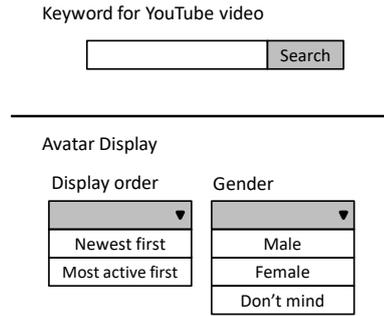


Figure 3: Search and selection panel

summary. Next, a desired video is selected from the list, and the screen changes to a WakWak Tube screen, consisting of the YouTube and avatar players shown in Figure 4, and playback starts automatically. At the same time, Kinect’s skeleton tracking begins and an avatar reflecting the user’s motion is displayed and recorded throughout the video.



Figure4: Screenshot of WakWak Tube

In Figure 4, the YouTube video player is on the top of the screen, and the avatars appear in the audience area at the bottom. The red avatar is the current viewer’s avatar, while the other colors, chosen from a color wheel, represent past viewers. Due to the limited space of the audience area, the number of past avatars is limited to eight. Collision detection is user to prevent the past avatars from overlapping.

Each avatar’s torso corresponds to the viewer’s gender. As shown in Figure 5, the male avatar’s torso (a) is a rectangle, and the female avatar’s (b) is a triangle.

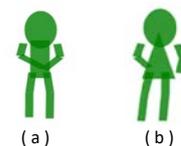


Figure 5: Male and female avatars

If the user’s own avatar does not have enough room to move in the audience area, the user is able to push other avatars away to make space (Figure 6).



Figure 6: Screenshot of the current user's avatar pushing away other avatars

4 IMPLEMENTATION

4.1 Device and software environment

We used Kinect for Windows SDK 2.0 for Kinect v2 to obtain the user's skeleton information. This sensor tracks as many as six complete skeletons consisting of 25 joints per person. WakWak Tube was developed using C# based on WPF, which is a GUI development library included in the Microsoft .Net framework.

4.2 Video search

Video search is based on keywords entered by the user in the text box in Figure 3. Passing a keyword string to the YouTube Data API and executing a search command returns a list of search results. This list includes information such as video ID, thumbnail URL, and caption. The video information acquired is displayed on the search result screen with a set of thumbnails and captions. A video ID tag is attached to each pair of thumbnail and caption.

When a thumbnail or caption detects a click event, the system transits to the movie playback screen. Dynamic creation of the video player is realized by passing its video ID tag as an argument of the playback screen class. The video player connects to the URL of the YouTube embedded player using the Web Browser control of WPF and displays the video.

4.3 Avatar information

Although Kinect v2 can acquire the skeleton coordinates of 25 joints of the human body, WakWak Tube utilizes the 14 points shown in Figure 7 to stabilize processing speed by suppressing the amount of skeleton information. After skeleton coordinates are acquired at 10 fps, they are converted to 2D coordinates for screen output and stored in an array. The avatar is represented by drawing lines from the shoulder and hip joints to the end of the fingertips, ellipses centering at the coordinate positions of the head, and rectangles or triangles centering at that of the spine.

4.4 Collision detection

Collision detection between avatars is carried out using the Intersect function of the Rect class, which is a function that returns the overlapping areas of two rectangles. Collisions between avatars are avoided by checking for overlapping portions and adding/subtracting an offset value corresponding to the overlap size to the avatar position

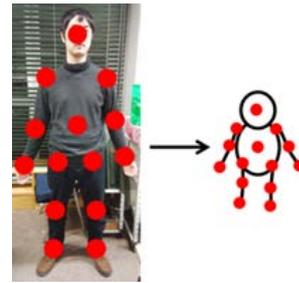


Figure 7: Joints used in WakWak Tube

coordinates. The rectangle used for collision detection is installed at the hand of the current viewer's avatar for pushing other avatars, and in the torso for approaching them. The torso of past avatars is used in both cases. The avatar position shift by the offset value is temporary, and the offset value set by a movement due to collision avoidance is not recorded in the file.

4.5 Recording avatar information

The behavior of the avatar—that is, the skeleton coordinate of each frame—is stored in a file. The avatar information is recorded from immediately after displaying the video player screen until the end of video playback. Avatar information is recorded in the file "video ID.csv" which is generated for each video ID in CSV format, as shown in Figure 8.

	A	B	C	D	E
1	fnum3230				
2	id1	oX200	oY150		
3	0X64	0Y38	1X75	1Y60	2X86
4	0X64	0Y39	1X75	1Y60	2X86
5	0X64	0Y39	1X74	1Y60	2X85
6	0X64	0Y39	1X74	1Y61	2X84
7	0X64	0Y39	1X73	1Y61	2X85
2928	id2	oX350	oY150		
2929	0X64	0Y38	1X75	1Y60	2X86
2930	0X64	0Y39	1X75	1Y60	2X86
2931	0X64	0Y39	1X74	1Y60	2X85
2932	0X64	0Y39	1X74	1Y61	2X84
2933	0X64	0Y39	1X73	1Y61	2X85

Figure 8: Example of "video ID.csv"

The first line of the CSV file is a header in which the maximum number of frames that can be recorded when videos are played to the end is described. Avatar information is stored from the next line of the header. The first row contains the user login ID and offset value. The offset value is used to shift the display position of the avatar without rewriting the skeleton coordinates. From the next line, 14 skeleton joints with 2D coordinates are recorded. Since the 2D coordinate information is stored on one line per frame, the frame number can be known from the number of rows. Although the skeleton coordinates are acquired as real numbers with Kinect, they are recorded after being cast to integer values. The avatar information for one person consists of a user ID, an offset value of coordinates, and a skeleton coordinate sequence. Since new avatar information is added from the next line of stored avatar information for each video playback, lower rows of the file are newer information. Skeleton information is recorded 10 times per

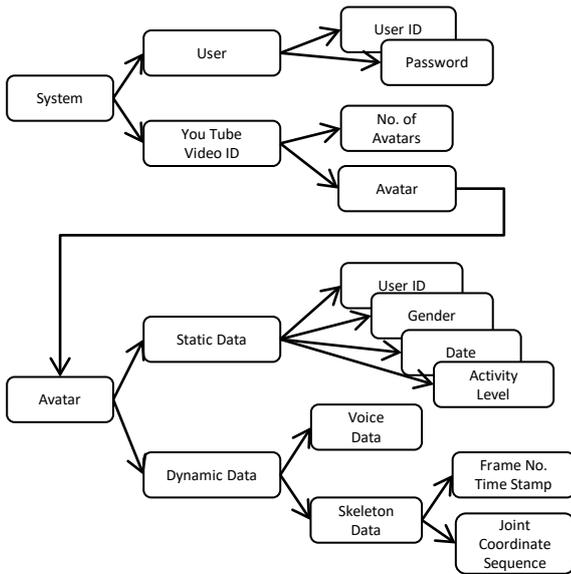


Figure 9: Data structure

second, the same as the avatar animation playback frame rate.

4.6 Data structure

Figure 9 shows the tree structure of the data managed by WakWak Tube. The system information consists of user login data and replayed YouTube video IDs, and the former consists of a user ID and password. YouTube video ID information has the number of avatars saved when a movie is being viewed; avatar data consists of static and dynamic data.

Static data include properties such as user name and gender, which are used when selecting avatars to display. Dynamic data consist of the sound data, frame number, timestamp, and skeleton coordinate sequence of each joint, and are used to play avatar animations in real time.

4.7 Static data format

Static data are represented in XML for search use and system extensibility, whereas the skeleton coordinate sequence in the dynamic data is stored in CSV format to improve its processing rate.

```
<? xml version="1.0" encoding="UTF-8" ?>
< StaticDataList >
  < StaticData avatarID="1" >
    < SelectionSetting >
      < Priority > newer </ Priority >
      < Gender > none </ Gender >
    </ SelectionSetting >
    < AvatarProperty userID = "0" >
      < DateTime > 2017/01/06 18:41:01 </ DateTime >
      < ActiveValue > 108.74 </ ActiveValue >
    </ AvatarProperty >
  </ StaticData >
</ StaticDataList >
```

```
</ StaticData >
< StaticData avatarID = "2" >
  <!-- ..... -->
</ StaticData >
</ StaticDataList >
```

Figure 10: Example of “StaticData.xml”

Figure 10 shows an example of the static data. Static data record selection history and link the avatar and user. The root element StaticDataList has StaticData, whose attribute is avatarID as a child element, and the child elements of StaticData are SelectionSetting and AvatarProperty. In SelectionSetting, the Priority and Gender elements specify the avatar display order and gender for selection processing. The AvatarProperty element, whose attribute is user ID has both a DateTime element that specifies the recording date and time and an ActiveValue element that has the activity level as a child element. Since we cannot calculate the ActiveValue until recording is complete, the file in Figure 10 is generated at the end of the recording.

4.8 LINQ to XML

We employ Language Integrated Query (LINQ), which provides an in-memory XML programming interface, to implement the selection function. Since database queries can be written in C# code, query processing is similar to SQL on not only XML files, but also arrays and lists in programs.

4.9 Calculation of avatar activity

The user of WakWak Tube can select an avatar to be played simultaneously with a YouTube video, for up to 8 avatars sorted in order of motion frequency. We now discuss a method for obtaining this activity indicator, from the skeleton coordinate sequence.

When the coordinate of SpineBase in frame i is (x_{is}, y_{is}) and the coordinate of Joint j is (x_{ij}, y_{ij}) , the distance $dist_{os}(i)$ between SpineBase and the origin and the distance $dist_{sj}(i)$ between each Joint and SpineBase can be obtained from the following formulas:

$$dist_{os}(i) = \sqrt{x_{is}^2 + y_{is}^2} \tag{1}$$

$$dist_{sj}(i,j) = \sqrt{(x_{ij} - x_{is})^2 + (y_{ij} - y_{is})^2} \tag{2}$$

In method A, we calculate the standard deviation $stdev(dist_{os}(i))$ of the coordinate values of the SpineBase, which is the center of the body, and $stdev(dist_{sj}(i,j))$, representing the variance in the distance between each Joint j and the SpinBase; the sum of the above values yields activity value A

$$A = stdev (dist_{os}(i)) + \sum_{j=1}^{14} stdev (dist_{sj}(i,j)) \quad (3)$$

(for $1 \leq i \leq maxframe$)

We obtain activity value B using the variance of the distance between the origin and the coordinates of each joint as follows:

$$dist_{oj}(i,j) = \sqrt{x_{ij}^2 + y_{ij}^2} \quad (4)$$

$$B = stdev (dist_{os}(j)) + \sum_{j=1}^{14} stdev (dist_{oj}(i,j))$$

(for $1 \leq i \leq maxframe$) (5)

In method C, $dif_{ij}(i,j)$ is the difference in the coordinates of Joint j between frame $i - 1$ and frame i .

$$dif_{ij}(i,j) = \sqrt{(x_{ij} - x_{i-1j})^2 + (y_{ij} - y_{i-1j})^2} \quad (6)$$

We sum $dif_{ij}(i,j)$ for each joint to find the whole difference $dif_i(i)$ in frame i .

$$dif_i(i) = \sum_{j=1}^{14} \sqrt{(x_{ij} - x_{i-1j})^2 + (y_{ij} - y_{i-1j})^2} \quad (7)$$

The average of the difference values for the entire frame is activity value C.

$$C = average (dif_i(i)) \quad (for \ 2 \leq i \leq maxframe) \quad (8)$$

The above three methods can be summarized as follows.

(A) standard deviation of the distance between the Spine-Base and each skeleton point; (B) standard deviation of screen coordinates of each skeleton point; (C) inter-frame movement distance of each skeleton point.

5 EXPERIMENT FOR ACTIVITY VALUE

We conducted experiments to examine whether the activity values obtained by methods A, B, and C coincide with subjective evaluation using avatar animations recorded with the following YouTube videos (Figure 11).



Figure 11: Videos used in experiments

- (1) Radio exercise video where three female instructors are showing body motion. (6 male and 3 female avatars are recorded with this video).

- (2) Miku Hatsune live concert video where the audiences are waving their hands (7 male and 2 female avatars)
 (3) Football goal scene collection video (6 male and 2 female avatars)

Seven men and one woman in their 20s were shown the recorded avatar animation and asked to rank the avatars' activity. In this experiment, all avatars were displayed as male, and the YouTube video was hidden.

We obtained the Spearman's rank correlation coefficients shown in Table 1 from the three algorithm rankings and the subjective evaluations.

Table 1 Spearman's rank correlation coefficient between algorithms and subjective evaluation

	Video 1	Video 2	Video 3
Activity A	0.62	0.26	0.7
Activity B	0.41	0.6	0.55
Activity C	0.54	0.7	0.81

In Video 1, viewers moved their whole body largely along with the video, so the change in distance increased and the rank correlation of Activity A increased. Since the actions in the video repeat, the standard deviation of the change in coordinates was small and the rank correlation of B was low.

When the audience watched the live concert video (Video 2), there was little change in skeletal distances, as viewers mostly just waved their arms. However, fast arm swings increased the difference between frames. As a result, Spearman's rank correlation was low in Activity A and high in Activity C.

In Video 3, since there were some avatars with big reactions (such as moving around to the right and left) to express pleasure, many changes in distance occurred frequently. As in the case of the live concert video, the difference between frames became large due to quick hand motions. As a result, the rank correlation increased in Activity A and C.

From the above observations, we determined that Activity A focuses on the distance between skeletons, and is suitable for whole-body movement, whereas Activity C focuses on frame difference and is suitable for quick motions.

6 CONCLUSION AND FUTURE WORK

WakWak Tube creates a virtual space in which a viewer's alter ego—in the form of an avatar—is displayed with the avatars of past viewers and added to video content. We examined feasible interactions to foster viewers' sense of connection with other avatars when the viewer's own avatar reflects his or her real-time motion and past viewers' avatars gather in the same audience space.

We implemented interaction functions between avatars, such as pushing away other avatars, and selection functions that enable a user to select the desired avatars to be displayed with his or her own avatar. The selection criteria

are date, gender, and activity level, the latter of which is calculated using a sequence of skeleton coordinates.

We examined three algorithms for obtaining activity values by three different types of videos, and found that algorithms using the difference between frames (Activity C) and the standard deviation of spine-based (Activity A) are superior to the algorithm using the variance of absolute coordinates (Activity B).

In future work, we believe that a more attractive system can be realized by adding the following functions:

- (1) Recording function: by recording and playing cheers of audiences, we expect that the user's sense of presence and unity will be further enhanced.
- (2) Notification function: designed to notify a user that his or her avatar was chosen by other users, which can be used as a trigger for communication via existing communication tools such as email or SNS tools.

In addition to the above, we must investigate the spatial structures needed to display more avatars in the audience area around YouTube video player.

REFERENCES

- [1] L. Oehlberg, N. Ducheneaut, J. D. Thornton, R.J. Moore, and E. Nickell, Social TV: Designing for distributed, sociable television viewing, Proceedings of the European Conference on Interactive TV'06, pp.251–259 (2006).
- [2] K. Chorianopoulos, Content-enriched communication–Supporting the social uses of TV, Journal of the Communications Network, Vol. 6, Part 1, pp.23–30 (2007).
- [3] C. Harrison and B. Amento, CollaboraTV: Using asynchronous communication to make TV social again; Interactive TV: A shared experience, Adjunct Proceedings of EuroTV, pp.218–222 (2007).
- [4] <https://www.rabb.it/> (accessed on June 9, 2017).
- [5] T. Shirokura, N. Munenaga, and T. Ono, ExciTube: Viewer's excitement sharing video player, Transactions of the VR Society of Japan, Vol. 18, No. 3, pp.247–254 (2013). (In Japanese)
- [6] A. Yoshida and H. Miyashita, Video sharing system that overlays body movements for the sense of unity, Proceedings of IPSJ Interaction 2012, pp.527–532, (2012). (In Japanese)
- [7] R. Vatavu, Audience silhouettes: Peripheral awareness of synchronous audience kinesics for social television, Proceedings of TVX'15, pp.13-22, (2015).

A Design and Implementation of Distributed Internet Live Broadcasting Systems Enhanced by Cloud Computing Services

Satoru Matsumoto* Yoshimasa Ishi* Tomoki Yoshihisa*
Tomoya Kawakami** Yuuichi Teranishi***

*Cybermediacenter, Osaka University, Japan

** Graduate School of Information Science, Nara Institute of Science and Technology, Japan

*** National Institute of Information and Communications Technology, Japan

Abstract -Recently, Internet live broadcasting services such as USTREAM, Twitcasting are prevailing. For Internet live broadcasting, the broadcasters record the videos using the cameras on their laptops or smart phones (clients). The clients send the video data to the broadcast servers of Internet live broadcasting services and the servers distribute the videos to the viewers. To improve the qualities of the videos or show additional information, the broadcasters sometimes add video or audio effects to their recorded videos. However, the types of the effects are limited in order that the processing times to add the effects do not disturb real time live broadcastings. The processing powers of the machines provided by cloud computing services often change. Thus, by flexibly select the machine to use and flexibly designate the effect processing, the clients can further relief the limit of the types of the effects. In this paper, we propose an Internet live broadcasting system enhanced by cloud computing services called a different world broadcasting system. In our proposed system, the policy for using cloud computing services, i.e., the selection of the machine to use and the effect designations, are described by ECA rules to be able to flexibly change the policy.

Keywords: Streaming Delivery; Multimedia; Internet Live Broadcasting; Video on Demand; Different Worlds Broadcasting.

1 INTRODUCTION

In recent years, real-time video distribution (Internet live broadcasting) via the Internet such as USTREAM and TwitCasting have become popular. In live Internet broadcasting, broadcasters or viewers often add video effects for comfortable Internet live broadcasting. So, by using high processing power clients, the broadcasters can add various types of the effects and further improve the visibilities or show more additional information. On the other hand, recently we can use high processing power machines provided by cloud computing services such as Microsoft Azure, Amazon Web Services, Google Cloud Platforms. To use these high power machines to add the effects, the clients select the machine to use, send the recorded video data to the selected machine and designate the detail of the effects (type, parameters, etc.)

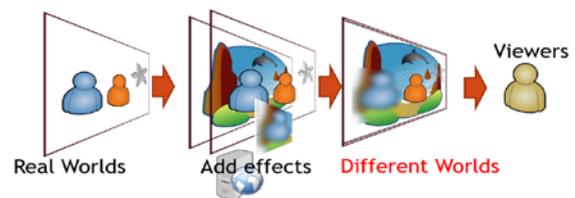


Figure 1. Different Worlds Broadcasting

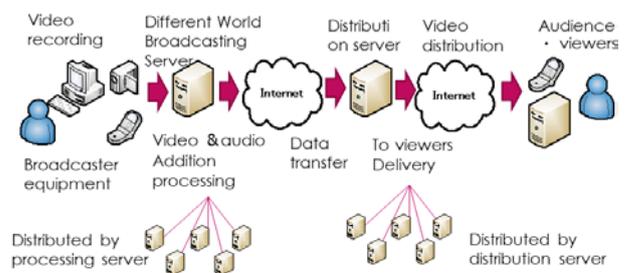


Figure 2. Overview of load distribution of effect addition processing

Although some distributed processing systems for Internet live broadcasting have been developed [1], these systems do not focus on cloud computing services. Our research group has developed a distributed Internet live broadcasting system called a different world broadcasting system. An image of different world broadcasting is shown in Figure 1. In different world broadcasting, effects such as blurring are added to the video shoot in the real world by a camera and distributed to viewers. In a different world broadcasting system, the different world broadcasting server performs the video effect. Therefore, the processing load caused by the addition of the video effect which was conventionally arose in the clients, are distributed (Figure 2). In different world broadcasting systems, it is assumed that a computer owned by a distributor or a virtual computer provided by a cloud service is used as a different world broadcasting server.

However, in our previous research, we have constructed a different world broadcasting system as using a computer owned by a distributor as a different world broadcasting server. Therefore, when many viewers request effect addition in a short period of time, or when some effects having large load are requested, it takes time to render the effects since the distributor's computer processes all of these requests. The delay by effect renderings makes the delivery frame rate unstable. For this reason, the available effects are

limited only to lightly loaded effects at present in order to prioritize the delivery frame rate. This limitation causes uncomfortable user experience of the different world.

On the other hand, in this research, we assume the use of a virtual computer on the cloud service as a different world broadcasting server which adds various effects to the video. In general, a plurality of virtual machines can be easily used in a cloud service. Therefore, by using multiple virtual machines as different world broadcasting servers, it is expected to realize high-speed effect addition while distributing the load among different world broadcasting servers and maintaining the delivery frame rate. However, in previous research, we assumed only use of a single computer. Therefore, in this research, we implement distributed Internet live broadcasting system using cloud service and evaluate its performance. In the implementation system, effect adding processing is performed using the virtual computer provided by the cloud service. When delegating processing to the virtual computer, it is possible to flexibly change the computer that performs processing using the ECA rule. As a result of the evaluation, we confirmed that the turnaround time of the effect adding process can be shortened by using the ECA rule.

The rest of the paper is organized as follows. We will introduce related research in Section 2. In section 3, we explain the design and implementation of a different world broadcasting system using cloud service and evaluate it in section 4. Finally, this paper is summarized in section 5.

2 RELATED RESEARCH

In MediaPaaS, for the purpose of a cloud computing service in the PaaS (Platform as a Service) format for live broadcasting, it is possible to encode, re-encode, distribute and image the video with the computer on the service provider side their propose a system [2].

Reference [3] implements a system that allows viewing (chasing playback) from a while ago when live broadcasting is performed using P2P network. In these video distribution systems, the point that broadcasters can not perform arbitrary video processing differs from the system proposed in this research. Several methods for reducing the delay time of video distribution have been proposed for Internet live broadcasting. In SmoothCache 2.0 [4], when live broadcasting, by caching video data from other peers and distributing it from the cached peer using the P2P network, it is possible to reduce the communication load and delay time of the video recording terminal. In Reference [5], that propose a distribution route determination method on a P2P network that minimizes the delay time. In the HD method, communication is reduced by transmitting video data to multiple viewing terminals at the same time using one-to-many broadcast distribution simultaneously with one-to-one communication [6]. Even in the proposed system, these delay time reduction methods can be applied at the time of video distribution, but this research is different from the video processing system. Further, researches on systems that perform video processing on stored video data are being conducted. In Reference [7], we propose a system that performs video processing by transferring video data shot by

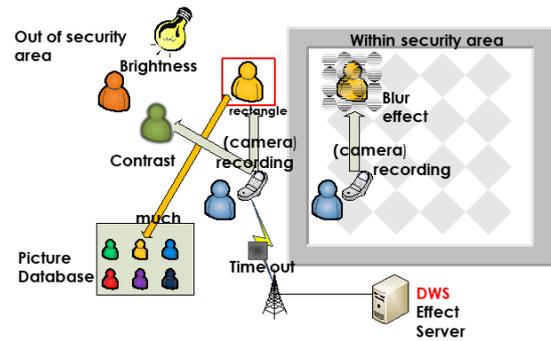


Figure 3. Events, conditions, Conceptual diagram of processing

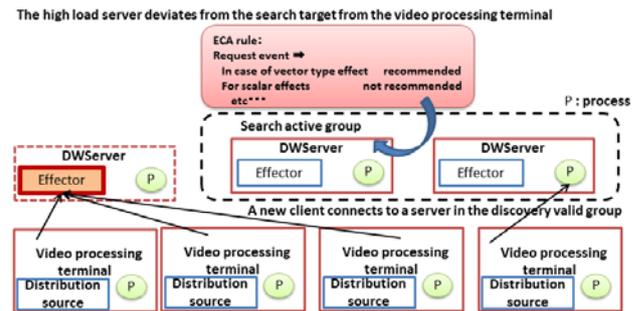


Figure 4. Applying ECA rules to load balancing Overview

a camera to a computer with high computing power. Reference [8] proposes a system that stores images captured by image capturing terminals with low computing power such as smartphones directly in an external storage device such as cloud storage without storing them in the image capturing terminal. These systems target stored video data and can not be used for Internet live broadcasting.

3 SYSTEM DESIGN AND IMPLEMENTATION

We applied a load distribution mechanism to different world broadcasting servers and video recording terminals of the different world broadcasting system.

This section shows the system design, the outline of the load distribution mechanism, the design of the description method used for dynamically allocating and executing the effect addition processing, and the implementation of these components.

3.1 Design Overview

In the different world broadcasting system using a cloud service, in order to distribute the load of virtual machines, it is necessary to describe processing to be assigned to each virtual machine. Kinds of processing are explicitly described and allocated when the kinds do not change frequently. When they frequently change, it is considered to dynamically allocate them according to kind of processing. In this research, processing is allocated by using ECA rules so that these process allocations can be flexibly described.

In previous design [1], in order to reduce the operation load of broadcasters, event-driven processing based on the

```

{
  "Rule A":{
    "eventname":"Set_effect",
    "condition":{
      "name":"Num_Find_Object",
      "object":"http://object1_haar.xml",
      "Value":">=1"
    },
    "action":{
      "name":"REQ_IP",
      "IP_address":"xx.xx.xx.xx"
    }
  },
  "Rule B":{
    "eventname":"Set_effect",
    "condition":{
      "name":"Blur"
    },
    "action":{
      "name":"REQ_IP",
      "IP_address":"yy.yy.yy.yy"
    }
  }
}

```

Figure 5. Early inquiries destination ECA rule for each effect

ECA (Event, Condition, Action) rule is automatically performed. The rule automatically starts by various events at the time of shooting and have the different world broadcasting servers add visual and sound effects.

For example, to protect portrait rights, the face authentication automatically detects specific persons in a database and add mosaics on the faces except the specific persons when entering an area. Also the system can automatically adjust the brightness of the video or adjust the contrast according to the brightness of the lighting (Figure 3).

In the previous research, ECA rules were applied to the different world broadcasting server side mainly for the purpose of reducing the operation load of the broadcaster, as described above. In this research, we designed the load balancing rule by ECA rule so as to avoid new video processing request to different world broadcasting servers in heavy load state. The ECA rules are distributed from the ECA rule distribution server to each server (Figure 4).

3.2 Overview of Load Balancing Mechanism

In assumed environments, there are three types of terminals or servers: video recording terminal, different world broadcasting server, video receiving terminal. The video recording terminal selects a different world broadcasting server capable of performing necessary video processing, and transmits the video effect library and the shot video. The different world broadcasting server operates on the virtual machine of the cloud service and performs the video processing on the video sent from the video recording terminal according to the instruction from the video recording terminal. The video processed by the different world broadcasting server is delivered to the video receiving terminal via the video distribution service. The video receiving terminal receives the processed video by selecting the server or the channel of the video distribution service which each handles the video to be.

In this research, focusing on the part where the video recording terminal selects a different world broadcasting server, we realize the load distribution of the different world broadcasting server. In the existing system, load distribution

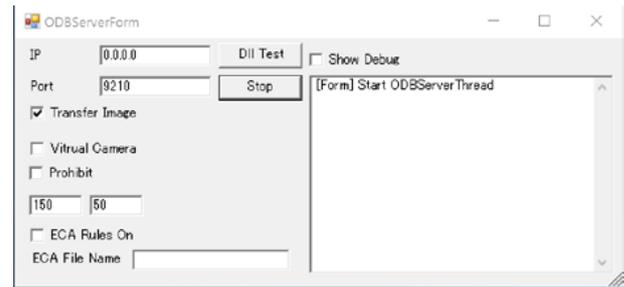


Figure 6. Screenshot of different world broadcasting server software

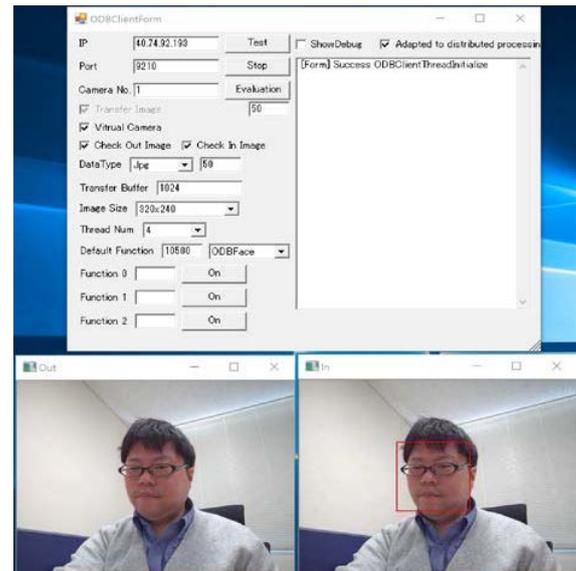


Figure 7. Screenshot of client software of video recording terminal

is realized by connecting the video recording terminal via the load distribution mechanism when connecting to the different world broadcasting server. In this method, since the load distribution mechanism selects a different world broadcasting server that can be used as a trigger for connection establishment from the video recording terminal and relays the communication there between, the video recording terminal and the different world broadcasting server cooperatively select the load distribution mechanism that can be used without being conscious of the existence of the mechanism. On the other hand, in the case of switching the different world broadcasting server while the video is being transmitted, for example when switching to a different world broadcasting server suitable for the scene or the effect, the video recording terminal disconnects the existing connection, that cannot switch the different world broadcasting server unless you reconnect. For this reason, it is difficult to switch the different world broadcasting server while smoothly continuing the video distribution.

In the different world broadcasting server, a different world broadcasting server software and a sever-side PIAX process are installed. PIAX [9] is Java based platform middleware that realizes serverless and efficient resource searching by utilizing the search function of the overlay

network, and it is provided as an open source software. The PIAX process in the video recording terminal and the different world broadcasting server are mutually connected via their overlay network.

The client-side PIAX process on the video recording terminal searches the overlay network in accordance with the request of the video recording terminal software, selects one different world broadcasting server that can be used from a plurality of different world broadcasting servers, returns a IP address of the different world broadcasting server and a listen port number of the different world broadcasting server software. Based on the response, the video recording terminal software establishes a connection with the different world broadcasting server software and starts sending the video. JSON-RPC is used for communication between video recording terminal software and client-side PIAX process, but since generally used JSON-RPC over HTTP is excessive quality, this system communicate JSON-RPC directly over TCP.

The server-side PIAX process on the different world broadcasting server controls its searching state according to the instruction from the different world broadcasting server software. The different world broadcasting server software that judged to be in a high load state himself connected to the control port of server-side PIAX process on the localhost and sends a leave request command from the search. As a result, the different world broadcasting server is excluded from the retrieval from the video recording terminal, and there are no more connections beyond the present. When the load of the different world broadcasting server software lowers and the state becomes ready to accept the new connection, it connects again to the control port of the server-side PIAX process on the localhost and sends a release leave request command from the search. Upon receiving this command, the server-side PIAX process again enters a state of accepting a search from the video recording terminal.

3.3 Designing ECA Rules

With the load distribution mechanism described in Section 3.2, it is possible to avoid new processing requests for different world broadcasting server which are under high load without constructing a load distribution system beforehand. Furthermore, on the video recording terminal side, by improving efficiency at first querying to a different world broadcasting server, we considered that time to completion of processing could be shortened. Depending on the type of video effect to be added, there may be cases where it is performed only on a server that has special processing calculation library software. In order to smoothly solve the initial value problem of this inquiry, we also use the ECA rule in the video recording terminal to solve this problem.

The event of ECA rule for load distribution is the effect selection on a video recording terminal software by user, and the rule has a list of initial inquiries IP address in advance. As a result, the server selection is automatically and smoothly performed also in the special video effect adding process. Figure 5 shows an example of the ECA rule.

In this example, the relation between the state and the action in the case where the video effect is selected by the pull-down menu as the event is described. The rule performs in case of the face recognition or blurring processing is selected as a condition, then it calls a function for obtaining an IP address as an action.

By describing this rule in advance, the server search operation of the user is greatly reduced, and the video processing request is dynamically performed to the optimum server only by selecting the video effect. If this rule is used as an initial value and the ECA rule distribution server distributes it to video recording terminal software simultaneously or the video recording terminal software downloads the rule at the time of launching the client software, it is possible for the users to use the system with zero configuration only by initial setting by the administrator.

3.4 Implementation

We used Microsoft Azure as a cloud service to deploy the different world broadcast system. The different world broadcasting server works on virtual machines provided by the Azure service. Each virtual machine is logically connected by a Virtual Network (VNet) which is one of the services of Microsoft Azure. Figure 6 shows the user interface of the different world broadcasting server software undertaking the video effect adding process. Figure 7 shows the user interface of the video recording terminal software that distributes the video at the video recording terminal. It can visually recognize the result of applying the effects.

If the check box of ECA rule on in the dialog box of the different world broadcasting server software is turned on, even if the user of the video recording terminal does not operate anything, the image effect is automatically added according to the ECA rule on the side of the different world broadcasting server.

In the ECA rule on the video recording terminal software side, the IP address for the video processing request to the different world broadcasting server is held. The video recording terminal software uses the IP address corresponding to the effect selected from the Default Effect pull down menu to perform the video processing request to the different world broadcasting server as an action. At this time, if the Attached to Distributed processing check box in the video recording terminal software dialog box is set to On, the initial video processing request is sent to the different world broadcasting server via the load distribution mechanism based on the ECA rules held by the client software. The processing request send to a different world broadcasting server with a margin. In this case, the ECA rule is described to inquire the load distribution mechanism about the IP address of the low-load foreign world broadcast server.

4 EVALUATION

In order to confirm the effectiveness of the proposed method, we have evaluated using the implementation system built on the virtual machine of the Microsoft Azure service.

The evaluation method, the evaluation environment, and the

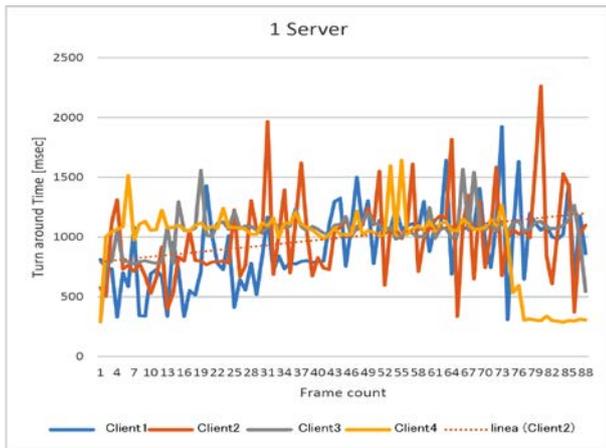


Figure 8. When there is one different world broadcasting server

evaluation result are below.

4.1 Evaluation Method

To evaluate the efficiency of our proposed system, we used Turnaround time including judgment time of ECA rule and video processing time as evaluation indexes. We compared and evaluated the turnaround time in the case where the video effect processing requests were concentrated from the viewer on the video effect processing server not using the ECA rule and the case where the video effect processing requests using the ECA rule were not concentrated. For the purpose of evaluation, we assigned multiple computers of the same performance to the cloud service. To select available different broadcast server, we used the overlay network PIAX described in the design section. When a specific server was overloaded from among the different world broadcasting servers, the server sent a notification to the sever-side PIAX process, and waits until the load is reduced. We measured the turnaround time in two cases. One was concentrated case which four video recording terminal softwares requested to one out of five different world broadcasting servers. The other was non concentrated case which each of the four video recording terminal softwares requested to a different world broadcasting servers-service we measured the time.

Since the main purpose of evaluation was video processing, live broadcasting software was not used in this evaluation. In the evaluation, a video processing that detects the face of a person in the video was used, assuming an example using the effect in the ECA rule. We increased the accuracy of detecting the human face every several frames and measured the turnaround time by applying calculation load. As a turnaround time, the time taken to transmit and receive frame data was used. For transmission and reception of frame data, preprocessing such as acquiring video from the camera from the previous reception completion until the start of next transmission, preprocessing such as acquiring video from the camera, and completion of transmission of frame data etc. from the video imaging terminal to the

different world broadcasting server , The time for

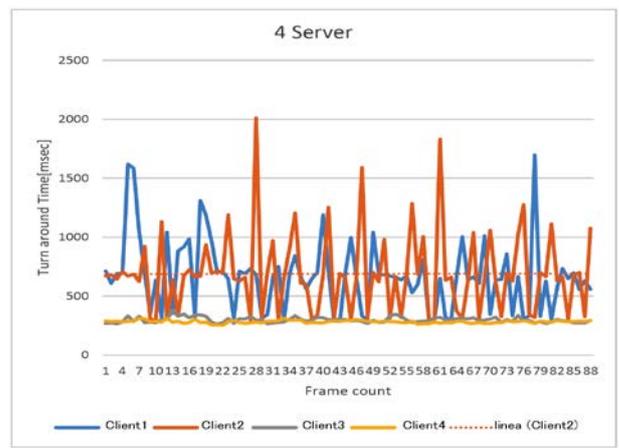


Figure 12. When there are four different world broadcasting servers

performing video processing on a different world broadcasting server, and the time from when the video recording terminal receives reception of frame data and the like from the different world broadcasting server. The video processing time was the time from the start of video processing upon receiving image data to the end thereof.

4.2 Evaluation Environment

The different world broadcasting server for evaluation worked on the virtual machines of the Microsoft Azure service. Table 1 shows the specification of the virtual machine and its OS. We used 5 different VMs for the different world broadcasting server. As a library that performs video processing on a different world broadcasting server, Open CV parallelized by Intel's parallel computing library TBB [10] was used.

Table1. Specification of Virtual Machine (different world broadcasting server)

Component	Performance
OS	Microsoft Windows Server 2016 Datacenter
Microsoft Azure (Virtual server service plan)	Standalone Server Microsoft Corporation Virtual Machine x64-based PC
CPU	Intel E5-2697 v3 Equivalent 2.4 GHz
Main memory	3,584 MB

4.3 Evaluation Results

Figures 8 and 9 show the evaluation results of the turnaround time in the evaluation environment described in section 4.2. Figure 8 shows the case of concentrated, and Figure 9 shows the case of non concentrated. The horizontal axis of the graph represents the frame count of shooting. The vertical axis of the graph represents the turnaround time in millisecond. In the scatter diagram of Figure 8 where the load concentrated on one different world broadcasting server,

the turnaround time monotonously increases, on the other hand, in Figure 9 where the video effect processing request was distributed with four different world broadcasting servers, the turnaround time increases in the latter half since the calculation load is increased in the latter half. However, at the first half frame count, there is no sudden increase in turnaround time as shown in Figure 8. In addition, although virtual machines with the same performance of the Azure service were used, the turnaround time varied that seems to be due to the network route to the different world broadcasting server and other causes. At the time of the first half frame count, the turnaround time in Figure 9 gently increases as compared with Figure 8. This indicates that the concentration of video processing is suppressed. We also measured the turnaround time to inquire about the recommended different world broadcasting server. The average value of one inquiry time of 50 trials was only 16.28 msec. It is performed only once at the beginning of the video processing, and does not affect the entire turnaround time.

As a result, it was confirmed that processing requests are allocated among different world broadcasting servers based on ECA rules, and the load can be distributed. Also, due to the effect of communication delay etc, it turned out that even if the hardware performance of the virtual machine is equal, the turnaround time may be interfered.

5 CONCLUSION

In this research, we implemented and evaluated a distributed processing type Internet live broadcasting (different world broadcasting) system using cloud services. By describing the processing assignment policy with the ECA rule, it is possible to flexibly change the virtual machine to be processed. In the implemented system, PIAX, which was used for searching and communicating with virtual machines. This makes possible to continuously allocate the video processing in the Internet live broadcasting even if the number of virtual machines increases or decreases. As a result of the evaluation, we confirmed that the turnaround time of the video processing can be shortened by using the proposed system.

In the future, we are planning to allocate processing in consideration of hardware performance of different world broadcasting servers, and to do minor processing on a video recording terminal.

ACKNOWLEDGMENT

This research was supported by Grand-in-Aids for Scientific Research (B) numbered 15H02702, for Challenging Exploratory Research numbered 26540045, And part of this research is the result of NICT · Osaka University joint research "research and development of advanced network platform technology for large scale distributed computing". We express our appreciation for writing it here.

REFERENCES

- [1] Satoru Matsumoto, Yoshimasa Ishi, Tomoki Yoshihisa, Tomoya Kawakami, and Yuuichi Teranishi, Different Worlds Broadcasting: A Distributed Internet Live Broadcasting System with Video and Audio Effects, in Proc. of 31st IEEE International Conference on Advanced Information Networking and Applications (AINA 2017), pp. 71-78 (2017).
- [2] B. Cheng, MediaPaaS: A cloud-based media processing platform for elastic live broadcasting, Proceedings of the 7th IEEE International Conference on Cloud Computing (CLOUD 2014), pp. 713-720 June (2014).
- [3] Y. Gotoh, T. Yoshihisa, H. Taniguchi, and M. Kanazawa, Brossom: a P2P streaming system for webcast, Journal of Networking Technology, Vol. 2, No. 4, pp. 169-181, Dec. (2011).
- [4] R. Roverso, R. Reale, S. El-Ansary, and S. Haridi, Smooth-Cache 2.0: CDN-quality adaptive HTTP live streaming on peer-to-peer overlays, Proceedings of the 6th ACM Multi-media Systems Conference (MMSys 2015), pp. 61-72, March (2015).
- [5] J. Dai, Z. Chang, and G.S.H. Chan, Delay optimization for multi-source multi-channel overlay live streaming, Proceedings of the IEEE International Conference on Communications (ICC 2015), pp. 6959-6964 June (2015).
- [6] T. Yoshihisa and S. Nishio, A division-based broadcasting method considering channel bandwidths for NVoD services, IEEE Transactions on Broadcasting, vol. 59, no.1, pp. 62-71 March (2013).
- [7] D. Gibbon and L. Begaja, Distributed processing for big data video analytics, IEEE ComSoc MMTC E-Letter, vol. 9, no.3, pp. 29-31 (2014).
- [8] W.-C. Ting, K.-H. Lu, C.-W. Lo, S.-H. Chang, and P.C. Liu, "Smart video hosting and processing platform for Internet-of-Things," Proceedings of the 2014 IEEE International Conference on Internet of Things (iThings 2014), pp. 169-176 Sept (2014).
- [9] Mikio Yoshida, Takeshi Okuda, Yuuichi Teranishi, Kaname Harumoto, Shinji Shimojyo: PIAX: A P2P Platform for Integration of Multi-overlay and Distributed Agent Mechanisms, Transactions of Information Processing Society of Japan / Information Processing Society of Japan, Vol. 49, No. 1, pp. 402-413 January (2008).
- [10] Thread Building Blocks, <https://www.threadingbuildingblocks.org/>

Session 5:
Position Recognition
(Chair: Yu Enokibori)

A Method for Estimating Indoor Position and Walking Speed by Using a Smartphone, with the Aim of Grasping People Flow in Indoor Passages

Masaru Matsubayashi* and Yoh Shiraishi**

*Graduate School of Systems Information Science, Future University Hakodate, Japan

**School of Systems Information Science, Future University Hakodate, Japan
{g2116044, siraisi}@fun.ac.jp

Abstract – The goal of our study is to grasp people flow by collecting the information of indoor position and walking speed of many pedestrians in indoor passages. In this paper, we propose a method for estimating the indoor position and walking speed of a pedestrian by using magnetic data and acceleration signal data. Each position has distinctive magnetic data due to the existence of ferromagnetic materials. We focus on the waveform of the magnetic data sequence that is measured while walking. We use these data sequences with distinguishing waveforms as location fingerprints, and estimate indoor position by comparing a targeted data sequence with the fingerprints. Also, the existing methods for walking speed estimation often generate the relationship between walking speed and acceleration signal features. However, there are differences between individuals in the relationship. Our method estimates walking speed by adjusting the relationship dynamically based on the indoor position estimation results. The experimental results showed that the error of indoor position estimation and walking speed estimation was approximately 1.36 m and 0.084 m/s, respectively.

Keywords: People flow; indoor positioning; walking speed estimation; magnetic field; smartphone

1 INTRODUCTION

The goal of our study is to grasp people flow in indoor passages to make people's movement smarter. We define people flow as the information of the velocity and volume of people passing through a certain passage per unit of time. The velocity of people flow decreases and the rate increases frequently because many people walk along indoor passages on a daily basis. We will be able to walk via unoccupied passages if we can grasp the people flow of these passages in advance.

There are a lot of conventional methods for measuring people flow by using cameras or laser range scanners [1-4]. These methods can grasp people flow correctly by counting and tracking people who are walking within the measurement area. However, in these methods the measurement area is restricted to the area around the installed equipment. In the case of grasping people flow in indoor passages throughout the whole of an indoor environment, we would need to install much more equipment in the indoor environment. These methods are inadequate for measuring people flow throughout all the indoor passages, because the cost of the equipment becomes high.

On the other hand, participatory sensing has been attract-

ing attention. Participatory sensing is an approach for grasping environmental conditions over a wide range by collecting sensor data and position information from many participants [5]. The data are measured using the participants' devices, such as smartphones. There are some existing methods of grasping crowd levels by participatory sensing [6,7]. Therefore, we consider this approach can be applied to grasping people flow in indoor passages.

The goal of our study is to grasp people flow by participatory sensing in indoor passages. In our approach, we collect indoor position and walking speed information from some pedestrian's smartphones. The information of indoor position and walking speed of a pedestrian is important for measuring people flow. If we can identify certain areas where some pedestrians cannot walk at a usual walking speed, we can estimate that people flow velocity in these areas is slow. Moreover, we can estimate crowd levels based on the findings discussed by Wirz *et al.* [8]. They show that the walking speed of pedestrians becomes slower as the crowd level becomes higher. In future work, we plan to estimate people flow by analyzing the statistics (e.g., mean, variance, maximum, minimum, etc.) of the walking speeds that are collected from many pedestrians.

In this paper, we describe a method to estimate indoor position and walking speed by using magnetic data and acceleration signal data that are measured by smartphones. We tackle a study on estimating indoor position and walking speed with the aim of grasping people flow in indoor passages. We have already proposed an estimation method by only using magnetic data measured by a smartphone as our previous work [9]. The difference between our method proposed in this paper and the previous work [9] is that our method uses acceleration signal data as well as magnetic data in order to capture short-term changes in walking speed.

Our method focuses on the waveform of the magnetic data sequence that is measured while walking. The waveform of the magnetic data sequence that is measured while a pedestrian is walking is different depends on the area through which the pedestrian has walked because each position has distinctive magnetic data due to the existence of ferromagnetic materials. Our method uses magnetic data sequences with distinguishing waveforms as location fingerprints, and estimates the area through which the pedestrian has walked (walked section). Also our previous method [9] estimates walking speed by dividing the length of the walked section by walking time. This method estimates the mean walking speed for 8 seconds (namely, the estimation interval is 8 seconds), and we cannot capture short-term changes in walking speed. In this paper, our method enables us to cap-

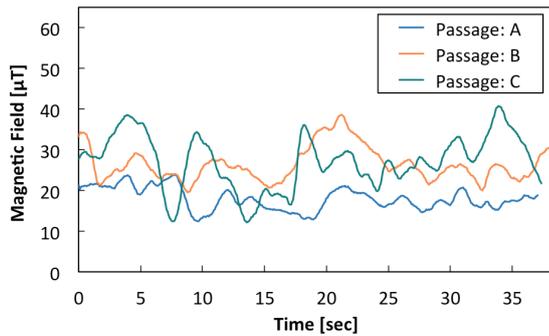


Figure 1: Magnetic data sequence measured in different passages

ture short-term changes in walking speed by using acceleration signal data as well as magnetic data that are measured by smartphones. The conventional methods for walking speed estimation by using acceleration signal features often generate the relationship between walking speed and acceleration signal features. However, there are differences between individuals in the relationship. Our method estimates walking speed by using the acceleration signal feature and by adjusting the relationship dynamically based on indoor position estimation results.

Preview of our experimental results. We evaluated our estimation method in an unoccupied passage and a crowded passage. As a result, in the indoor position estimation, the mean error of our method was approximately 1.36 m. In contrast, the mean error of a naive method was approximately 11.45 m. In the walking speed estimation, the mean error of our method was approximately 0.084 m/s, while the mean error of a conventional method was approximately 0.089 m/s.

Contributions. The main contributions of this paper are as follows:

- Our method enables us to capture short-term changes in walking speed, by using acceleration signal data as well as magnetic data.
- Our method estimates walking speed by adjusting the relationship between walking speed and acceleration signal feature dynamically, based on the indoor position estimation results. In our method it is not necessary to learn the relationship, and the method enables us to estimate the speed with accuracy equal to that of the existing methods.

2 RELATED WORK

In this section, we briefly discuss related works dealing with aspects such as crowd level estimation by participatory sensing, indoor position estimation, and walking speed estimation, in Sections 2.1, 2.2, and 2.3, respectively. Also, we summarize the related works in Section 2.4.

2.1 Crowd level estimation by participatory sensing

Kannan *et al.* [6] count the number of people by putting out an audio tone from a smartphone speaker and having the tone be picked up by other smartphones in the vicinity.

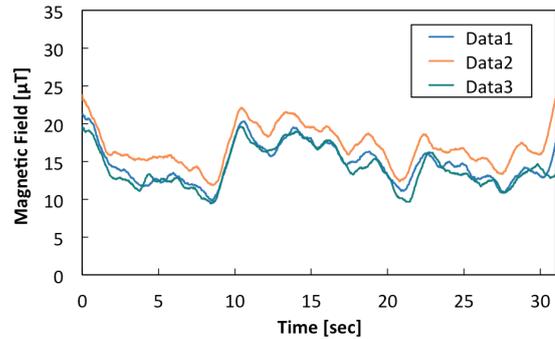


Figure 2: Magnetic data sequence measured in the same passage

Their method can count the number of people correctly if all pedestrians have smartphones on which their system is installed. However, their method does not estimate people flow velocity. Nishimura *et al.* [7] estimate crowd levels in indoor environments by using an accelerometer and microphone equipped on a smartphone to measure step-intervals and the volume of people's walking noise, respectively. Step-interval is the time between the current step and the next step. In future work of this study, the estimation result and the current position will be uploaded to a server after a pedestrian has estimated the crowd level of his or her current position. In Nishimura's study, they did not propose a method for estimating people flow velocity. In addition, consideration of the indoor position estimation method is insufficient. They will use the indoor positioning method based on Wi-Fi fingerprints. However, the accuracy of this method decreases in crowded environments because the radio waves of the Wi-Fi are attenuated by people's bodies.

2.2 Indoor position estimation

Indoor position estimation [10-12] has attracted a lot of interests, and there are many technologies for it such as using BLE beacons, Wi-Fi fingerprinting, magnetic fingerprinting, and so on. In our study, we need select an indoor positioning technology that can satisfy the following requirements for grasping people flow.

- Indoor position estimation without installing new equipment.
The goal of our study is to grasp people flow in indoor passages. We need a method with low cost to have many participants because we introduce an approach based on participatory sensing in order to archive the goal.
- Indoor position estimation with fine granularity.
We need grasp crowded spots such as an area in front of a door and a notice board.

As known methods that can estimate indoor position without installing new equipment, there are Wi-Fi fingerprint based methods and magnetic fingerprint based methods. However, as discussed in Section 2.1, the accuracy of the method based on Wi-Fi fingerprint decreases in crowded environments because the radio wave of the Wi-Fi is attenuated by the people's bodies. The methods based on Wi-Fi fingerprint are inadequate for our study because it is important to estimate crowded spots with fine granularity in

our study. Therefore, we focus on the positioning methods based on magnetic fingerprint, which can estimate indoor position with fine granularity even crowded environments.

Each position has distinctive magnetic data due to the presence of different kinds of pillars, doors, elevators, etc., that include ferromagnetic materials like steel or iron. Figure 1 shows the magnetic data sequences that are measured while a pedestrian is walking in three different passages. Figure 2 shows the magnetic data sequences that are measured while a pedestrian is walking in the same passage. The indoor positioning method based on magnetic fingerprint [13-15] employs a scene analysis technique. This method consists of a fingerprint generation and an estimation phase. The fingerprint generation phase is always executed before the estimation phase. In the fingerprint generation phase, magnetic data are collected at pre-selected indoor reference positions. For each reference position, magnetic data is obtained as a fingerprint. The fingerprints for all reference positions form a database. In the estimation phase, the position of the pedestrian is estimated based on the magnetic data that is obtained at the current position. The magnetic data of the current position is compared to all the fingerprints, and the reference position with the highest similarity is returned as the estimated position.

It should be noted that there is a difference between the magnetic data even if these are measured in the same position. As shown in Figure 2, the orange line of the magnetic data sequence is higher than others, even though these are measured in the same passage. The reason why we obtain this result is the differences in the situations, such as the height of the position at which the smartphone is held, the magnetization of iron around the magnetometer equipped on the smartphone, and the accuracy of the calibration of the magnetometer. In these situations, the accuracy of indoor position estimation decreases.

2.3 Walking speed estimation

The conventional methods for walking speed estimation [16-18] often generate the relationship between walking speed and acceleration signal features, and estimate walking speed based on this relationship. Kouroggi and Kurata [16] learn the relationship by using linear regression. Vathsangam *et al.* [17] learn the relationship by using Gaussian process regression. Park *et al.* [18] learn the relationship by using regularized least squares.

These methods have the problem of differences between individuals in the relationship. These conventional methods [16-18] solved this problem by generating the relationship for each pedestrian in advance. However, it is difficult to generate the relationship for a large number of pedestrians in advance.

2.4 Summary of related works

The indoor position estimation method based on magnetic fingerprint is known to be a method that can estimate indoor positions accurately. The accuracy does not decrease in crowded environments because the magnetic data is not attenuated by the human body, unlike the radio waves of the Wi-Fi. However, the methods based on magnetic data fin-

gerprint have the problem that there are differences between the magnetic data even if they are measured in the same position. This problem causes the decrease in accuracy.

The conventional methods for walking speed estimation by using the acceleration signal features have the problem of differences between individuals in the relationship between the walking speed and the features. If we can generate the relationship for each pedestrian in advance, we can estimate walking speed accurately. However, it is difficult to learn the relationship of a large number of pedestrians in advance.

3 METHOD

In this section, we describe our method in detail. We mention a proposal of our study, research tasks and approaches, an overview of our method, and the detailed procedure of our method in Sections 3.1, 3.2, 3.3, and the sections after Section 3.4, respectively.

3.1 Proposal of our study

In this paper, we propose a method for indoor position and walking speed estimation by using magnetic data and acceleration signal data. For estimating indoor position, we focus on a waveform of the magnetic data sequence that is measured while walking. Our method uses these data sequences with distinguishing waveforms as location fingerprints, and estimates indoor position by comparing a targeted data sequence with the fingerprints. For estimating walking speed, our method estimates by using the acceleration signal feature and by adjusting the relationship dynamically based on the position estimation results.

3.2 Research tasks and approaches

The main research tasks of this paper are as follows:

- **Research task 1:** To solve the problem that the accuracy of indoor position estimation decreases because of differences between magnetic data, even if the data are measured in the same position.
- **Research task 2:** To solve the problem that we need to generate the relationship between walking speed and acceleration signal feature for each pedestrian in advance in order to estimate walking speed accurately.

The approaches to our research tasks are as follows:

- **Approach to research task 1:** We focus on a waveform of the magnetic data sequence that is measured while walking. Waveforms of the magnetic data sequences measured in the same passage are similar to each other even if there is a difference between the magnetic data. Our method uses these data sequences with distinguishing waveforms as location fingerprints, and estimates the indoor position by adjusting an offset and comparing a targeted data sequence with the fingerprints.
- **Approach to research task 2:** Our method estimates walking speed by adjusting the relationship dynamically based on the position estimation results. In our method, we estimate walking speed by a simple linear model, as follows:

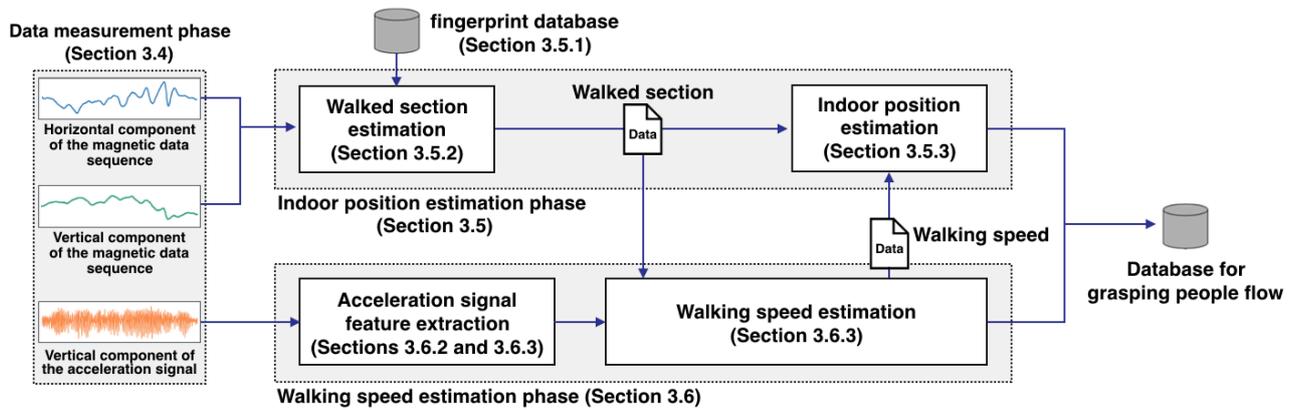


Figure 3: Overview of our method

$$v = kx \quad (1)$$

where v is the walking speed, x is the acceleration signal feature, and k is the constant value for converting to walking speed. k has the problem of differences between individuals. In our method, we set the proper k dynamically based on the walking distance calculated by the difference between two indoor position estimation results.

3.3 Overview of our method

Figure 3 shows the overview of our method. Our method consists of a data measurement phase, indoor position estimation phase, and walking speed estimation phase.

In the data measurement phase, while a pedestrian is walking we measure the magnetic data sequence and the acceleration signal data by using a magnetometer and an accelerometer equipped on a smartphone. We describe the detail of this phase in Section 3.4.

In the indoor position estimation phase, we use the data sequences with distinguishing waveforms as location fingerprints, and estimate the indoor position by comparing a targeted data sequence, which is measured while a pedestrian is walking for t seconds, with the fingerprints. Then, we use the horizontal and vertical components of the magnetic data sequence. As an estimation result, we get start and end positions of the area through which the pedestrian has walked for t seconds. We call the area between the start and end point the walked section. This estimation result will be used in indoor position estimation and walking speed estimation. After the walking speed has been estimated, we estimate the indoor position at each second by using the results of walking speed estimation, because the walked section corresponds to an area through which the pedestrian has walked for 30 seconds but does not correspond to a position. We describe the details of this phase in Section 3.5.

In the walking speed estimation phase, firstly we extract the acceleration signal feature that has the highest correlation with walking speed. Next, we adjust the k in Equation 1 by using the walking distance calculated by the difference between the start point and the end point of the walked section. Finally, we estimate the walking speed using Equation 1. We describe the details of this phase in Section 3.6.

3.4 Data measurement phase

In the data measurement phase, while a pedestrian is walking we measure the magnetic data sequence and the acceleration signal data by using a magnetometer and an accelerometer equipped on a smartphone. Then, the sampling rates of the magnetometer and the accelerometer are set to 50 Hz and 100 Hz, respectively.

Magnetic data is measured as three-dimensional vector quantities on three axes: the X-, Y-, and Z-axis. If the inclination of the smartphone (to what degree it is vertical or horizontal in relation to the floor) changes, magnetic data of the smartphone-based coordinate system changes. Therefore, we need to convert the magnetic data of the smartphone-based coordinate system into that of a world coordinate system. We can compute the Z-axis component of the magnetic data of the world coordinate system by detecting the direction of gravity. However, we cannot compute the X-axis and Y-axis components of the magnetic data of the world coordinate system, because it is difficult to detect the smartphone inclination due to the influence of disturbance of the magnetic field in indoor environments. Since our method is desired to be robust with respect to the inclination of the smartphone, we compute the vector quantity of the vertical and horizontal components of the magnetic data. The horizontal component of the magnetic data is the vector norm of a component obtained by composing the X-axis and Y-axis components of the magnetic data of the world coordinate system.

After we measure the magnetic data sequence, we smooth the magnetic data sequence with the simple moving average method. Finally, we extract input data sequences from the magnetic data sequence and acceleration signal data by partitioning them with a sliding-window of 30 seconds (The above mentioned parameter t equals 30).

3.5 Indoor position estimation phase

In this section, we describe the stage of the indoor position estimation phase. These stages are generation of fingerprints, walked section estimation, and indoor position estimation, described in Section 3.5.1, 3.5.2, and 3.5.3, respectively.

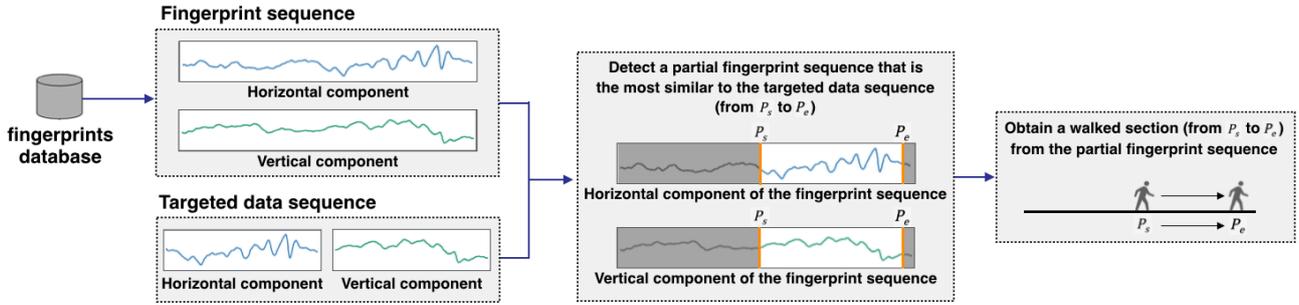


Figure 4: Overview of the walked section estimation

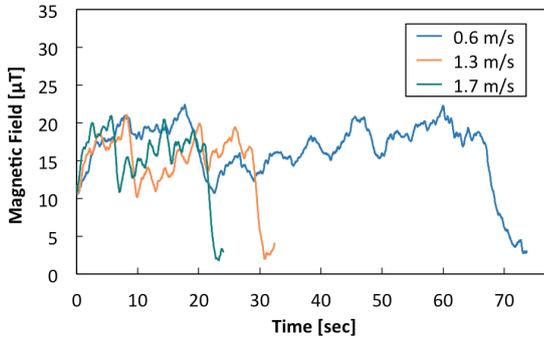


Figure 5: Magnetic data sequences measured by a pedestrian walking at three constant speeds (0.6 m/s, 1.3 m/s, and 1.7 m/s) in the same passage.

3.5.1. Generation of fingerprints

We represent the targeted indoor environment with passages as a graph structure that consists of node and links. Each crossroad of passages is represented as a node and each passage is represented as a link. In this paper, we assume that each passage is a straight line without corners. In each passage, we set reference lines in parallel to the passage, because most people walk straight in daily life. We measure a magnetic data sequence by walking along a reference line in each passage at a constant speed. Finally, we generate fingerprints by relating each item of magnetic data to the corresponding position, and these generated fingerprints are recorded in a fingerprint database.

3.5.2. Walked section estimation

Figure 4 shows the overview of the walked section estimation. In the walked section estimation, we estimate the walked section that a pedestrian has walked for 30 seconds. In advance, a targeted data sequence was extracted as described in Section 3.4. Firstly, we compare the waveforms of the targeted data sequence with all fingerprints and detect a partial fingerprint sequence that is the most similar to the targeted data sequence. At that time, we have a problem in that the magnetic data sequence is stretched and shrunken along the time axis as the walking speed change. Figure 5 shows the magnetic data sequences when a pedestrian walked in the same passage at different constant speeds. In this case, if we use a simple sequence matching algorithm such as Euclidean distance, we cannot recognize the mag-

netic data sequences shown in Figure 5 as sequences measured in the same passage. Therefore, we use SPRING [19] to detect the partial fingerprint sequence that is the most similar to the targeted data sequence. SPRING is a method for detecting subsequences that are similar to a given query sequence from data streams, by stretching and shrinking the sequence along the time axis to minimize the distance between the two sequences. Finally we obtain a walked section from the partial fingerprint sequence that is the most similar to the targeted data sequence.

3.5.3. Indoor position estimation

Indoor position estimation is always executed after walking speed estimation. We cannot estimate indoor position at each second only to estimate the walked section because we estimate the walked section as the area through which the pedestrian has walked for 30 seconds. Also, the estimated section does not correspond to a point. Firstly, we estimate walking speed at each second as described in Section 3.6.3. Secondly, we estimate indoor position at each second by cumulatively adding the distance walked in each second to the start point of the walked section.

3.6 Walking speed estimation phase

In this section, we describe the walking speed estimation phase. The phase comprises selection of the estimation model, selection of the acceleration signal feature, and walking speed estimation, described in Sections 3.6.1, 3.6.2, and 3.6.3, respectively.

3.6.1. Selection of the estimation model

In the conventional methods for walking speed estimation, walking speed is estimated by using polynomial models. These methods can estimate walking speed correctly by learning or inputting the proper parameters for each pedestrian in advance. However, it is difficult to learn and input the proper parameters for a large number of pedestrians. Therefore, we estimate walking speed by adjusting the proper parameters dynamically. In order to do this, we use the simple linear model (Equation 1), as shown in Section 3.2. One of the advantages of this model is that it can adjust the parameter easily because the number of parameters is one. In our method, we use the walking distance calculated by the difference between the start point and the end point of the walked section. If the number of parameters is more than

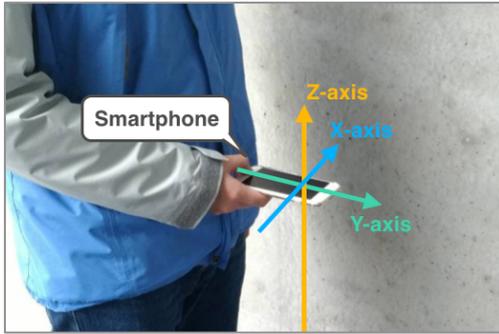


Figure 6: How to carry a smartphone

one, we need a lot of information to adjust parameters. However, we consider that, other than walking distance, we cannot obtain effective information to adjust parameters dynamically. Another advantage is that there are a lot of acceleration signal features that have high correlation with walking speed. Therefore, we use the simple linear model (Equation 1) shown in Section 3.2.

3.6.2. Selection of acceleration signal feature

We need to select the acceleration signal feature that has the highest correlation with walking speed if we use the simple linear model (Equation 1) shown in Section 3.2. In our method, we select standard deviation as the feature, based on the result of the preliminary experiment to select the acceleration signal feature that has the highest correlation with walking speed.

In the preliminary experiment, firstly, we collected the vertical component of the acceleration signal data from five participants walking along a 50 m passage. The participants consisted of five men of varying heights. Each participant was asked to walk along the passage at six different speeds (0.4, 0.6, 0.8, 1.0, 1.2, and 1.4 m/s) for three minutes. When the participants walked, they were asked to carry the smartphones in front of their bodies, as the Z-axis of the smartphone is always set to the vertical direction, as shown in the Figure 6. Then, the sampling rates of the accelerometer are set to 100 Hz. Next, we partitioned the vertical component of the acceleration signal data with a sliding-window of 512 samples (5.12 seconds) with a 50% overlap between subsequent windows, and we calculated the five features of each piece of the vertical component of the acceleration signal data that was partitioned. The five features are mean, maximum, minimum, standard deviation, and the sum of the squared magnitudes of the components of the signal that was used in the method of Park *et al.* [18]. We calculated the correlation between these features and walking speed.

As a result, standard deviation has the highest correlation with walking speed ($R = 0.75$ in our data). The sum of the squared magnitudes of the components of the signal has weaker correlation ($R = 0.71$ in our data) than the correlation between standard deviation and walking speed. Therefore, we select standard deviation as the acceleration signal feature to estimate walking speed.

3.6.3. Walking speed estimation

In walking speed estimation, firstly we partitioned the ac-



(a) The unoccupied passage (b) The crowded passage

Figure 7: Experimental passages

celeration signal data, that was measured while a pedestrian was walking, with a sliding-window of 400 samples (4 seconds) with 75% overlap between subsequent windows, and we calculated the standard deviation (SD_i) of each section of acceleration signal data that was partitioned. At that time, we needed to set the optimal window size. If window size is set larger than the optimal window size, we cannot capture short-term changes in walking speed. On the other hand, if window size is set smaller than the optimal window size, the accuracy of walking speed estimation decreases. In our method, we set 400 samples as the optimal window size, based on the result of the preliminary experiment to set the optimal window size. After we calculated the SD_i , we estimated walking speed by

$$v_i = \frac{\bar{v}}{SD} \cdot SD_i \quad (2)$$

where v_i is walking speed, \bar{v} is the mean walking speed calculated by dividing walking distance in the walked section by 30 seconds, and \overline{SD} is the mean of SD_i .

4 EVALUATION

In this section, we describe the evaluation of our method, which comprises the experimental setup and results, described in Sections 4.1 and 4.2, respectively.

4.1 Experimental setup

We conducted an experiment to evaluate our method in the passages shown in Figure 7. The length of the passage is 50 m, and the width is 2.5 m. The participants consisted of four men of different heights, who were labeled participant A, B, C, and D. In the experiment, we evaluated our method in two scenarios as follows:

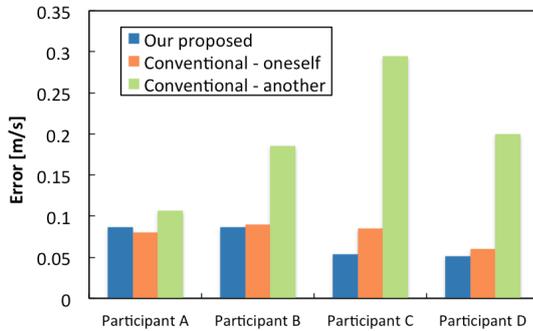
- **Scenario 1:** Freely walking in the unoccupied indoor passage shown in Figure 7(a).
- **Scenario 2:** Freely walking in the crowded indoor passage shown in Figure 7(b).

When the participants walked in these passages, they were asked to carry the smartphones in front of their bodies, as the Z-axis of the smartphone is always set to the vertical direction, as shown in Figure 6.

As an advance preparation for the experiment, we generated four fingerprints in the experimental passage and three passages around the experimental passage on the same floor.

Table 1: The errors (in meters) of indoor position estimation in scenario 1

Participant	Naive method		Proposed method	
	Mean	Median	Mean	Median
A	0.68	0.46	1.40	1.16
B	25.33	20.27	8.37	0.69
C	2.78	3.05	1.89	1.59
D	25.18	26.03	0.63	0.49

**Figure 8: The errors of walking speed estimation in scenario 1**

When we collected magnetic data sequences to generate the fingerprints, we set a reference line in the center of the passage, and the participant walked along the reference line at a constant speed of 1.3 m/s.

We evaluated the accuracies of indoor position and walking speed estimations. In the evaluation of indoor position estimation, we compared our method with a naive method. In the naive method, we give the horizontal and vertical components of the magnetic data sequence as a targeted data sequence. The naive method estimates indoor position by comparing the targeted data sequence with all the fingerprint sequences and detecting the partial fingerprint sequence that is the most similar to the targeted data sequence by not adjusting the offset.

In the evaluation of walking speed estimation accuracy, we compared the accuracies of our method and a conventional method proposed by Park *et al.* [18]. In the conventional method, we evaluated the accuracy in two cases: using an estimation model learned from one's own data (*Conventional method - oneself*), and using an estimation model learned from the data of another participant, E, whose height is the mean height of our participants (*Conventional method - another*).

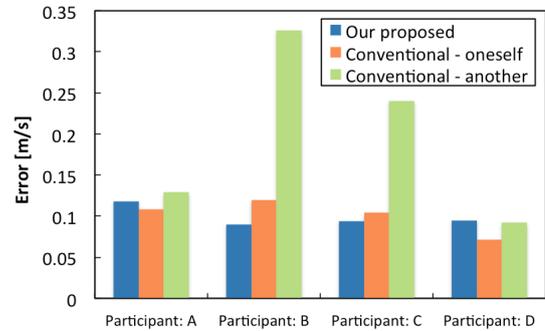
We measured correct indoor position and walking speed by relying on colored tiles that occurred every 0.5 m in our test passage. When the participant walked along the test passage, video of the participant was captured from diagonally above. We calculated the correct indoor position and walking speed based on the video.

4.2 Results

In this section, we describe the results of the indoor position estimation and the walking speed estimation in Section 4.2.1 and 4.2.2, respectively.

Table 2: The errors (in meters) of indoor position estimation in scenario 2

Participant	Naive method		Proposed method	
	Mean	Median	Mean	Median
A	8.71	6.01	10.54	1.79
B	7.69	4.34	12.18	2.55
C	4.75	4.90	5.38	1.23
D	26.36	26.53	11.03	1.44

**Figure 9: The errors of walking speed estimation in scenario 2**

4.2.1. Indoor position estimation

We describe the results of indoor position estimation. We used not only the mean of the error but also the median of the error as an evaluation index, because there is a possibility that a position very far from the correct position will be estimated in rare cases. If such a position is estimated, the mean error becomes large. In this case, we cannot evaluate correctly just by using the mean of the errors. Therefore, we used not only the mean of the error but also the median of the error as an evaluation index.

Table 1 and Table 2 show the mean and median errors of the indoor position estimation performed by our method and the naive method for every participant. Table 1 shows the estimation errors in scenario 1. Table 2 shows the estimation errors in scenario 2.

Result 1. As shown in Table 1, the mean and median errors estimated by the naive method were smaller than our method in the case of participant A walking in the unoccupied passage.

Result 2. As shown in Table 2, in the case of all the participants other than D walking in the crowded passage, the mean errors estimated by the naive method were smaller than our method. At the same time, the median errors estimated by our method were smaller than those of the naive method.

Result 3. As shown in Table 1 and Table 2, most of the mean and median errors in scenario 1 were smaller than in scenario 2.

Result 4. As shown in Table 1, the mean and median of the errors estimated by our method were smaller than those of the naive method, except in the case of participant A walking in the unoccupied passage.

4.2.2. Walking speed estimation

We describe the results of walking speed estimation. We use mean of the error as an evaluation index. As mentioned in Section 4.2.1, in our method for indoor position estimation, there is a possibility that a position very far from the correct position will be estimated in rare cases. If such a position is estimated, it acts as noise data when used to grasp people flow. However, we can remove these data easily by considering the front-rear relationship between the estimation results because the mixing ratio of these data is less than twenty percent. Therefore, we remove the noise data when we evaluate the accuracy of walking speed estimation.

Figure 8 and Figure 9 show the mean errors of the walking speeds estimated by our method, *Conventional method - oneself*, and *Conventional method - another* for every participant. The vertical axis corresponds to mean estimation error. Figure 8 and Figure 9 show the estimation errors in scenario 1 and scenario 2, respectively.

Result 5. As shown in Figure 8 and Figure 9, the mean errors estimated by *Conventional method - another* were frequently larger than those of the other methods. The means of all estimation errors were 0.196 m/s and 0.198 m/s in scenario 1 and scenario 2, respectively. On the other hand, our method and *Conventional method - oneself* did not show such a result. In scenario 1, the mean errors of all participants estimated by our method and *Conventional method - oneself* were 0.069 m/s and 0.078 m/s, respectively. In scenario 2, the mean errors of all participants estimated by our method and the *Conventional method - oneself* were 0.099 m/s and 0.100 m/s, respectively.

5 DISCUSSION

In this section, we describe the indoor position estimation results and walking speed estimation results in Section 5.1 and 5.2, respectively.

5.1 Indoor position estimation

As show in Result 1, the mean and median errors of the indoor position estimated by the naive method were smaller than our method in the case of participant A walking in the unoccupied passage. This is because the strength of the targeted data sequence was very similar to the correct fingerprint sequence. In our method, the information of strength is lost because we adjust the offset between the targeted data sequence and the fingerprints. Therefore, if the strength of the targeted data sequence is very similar to the correct fingerprint sequence, the naive method can estimate more accurately more than our method.

As shown in Result 2, in the case of all the participants other than D walking in the crowded passage, the mean errors of the indoor position estimated by naive method were smaller than those of our method. This is because the number of positions estimated by our method that were very far from the correct position was larger than that of the naive method. In our method, we obtained this result more often than the naive method because the information of strength is lost, as mentioned above. On the other hand, as shown in

Result 2, the median errors of the indoor position estimated by our method were smaller than those of the naive method. This is because the number of positions estimated by our method that were very near to the correct position was larger than that of the naive method.

As show in Result 3, most of the mean and median errors of indoor position estimation in scenario 1 were smaller than in scenario 2. In scenario 1, participants had walked straight. On the other hand, in the crowded passage, participants had not walked straight, in order to avoid other people. In our method, we set a straight reference line in the center of the passage. In this case, the similarity between the targeted data sequence and correct fingerprint sequence decreased because there is a difference between the magnetic data sequences measured at the walls side and center of a passage. Therefore, the mean and median errors in scenario 1 were smaller than the errors in scenario 2.

As show in Result 4, the mean and median errors of the indoor position estimated by our method were smaller than those of the naive method, except in the case of participant A walking in the unoccupied passage. This is because there is a difference between the strengths of the targeted data sequence and correct fingerprint sequence. This result shows that our method can solve the problem that the accuracy of indoor position estimation decreases because of differences between magnetic data even when measured in the same place.

5.2 Walking speed estimation

As shown in Result 5, the mean errors of the walking speed estimated by the *Conventional method - another* were frequently larger than those of the other methods. On the other hand, our method and the *Conventional method - oneself* did not show such a result. This result shows that our method can solve the problem that we need to generate the relationship between the walking speed and the acceleration signal feature for each pedestrian in advance in order to estimate walking speed accurately.

6 CONCLUSION

The goal of our study is to grasp people flow by collecting indoor position and walking speed information from some pedestrians in indoor passages. In this paper, we proposed a method for indoor position and walking speed estimation by using a magnetometer and accelerometer equipped on a smartphone. We evaluated our method in an unoccupied passage and a crowded passage. As a result, the mean error of indoor position estimation was approximately 1.36 m, while the mean error of walking speed estimation was approximately 0.084 m/s.

Contributions. The main contributions of this paper are as follows:

- Our method enables us to capture short-term changes in walking speed, by using acceleration signal data as well as magnetic data.
- Our method does not require learning of the relationship between walking speed and acceleration signal features, and enables us to estimate the speed with accuracy equal to that of the existing methods.

In future work, we plan to improve our method to enable highly accurate estimation of indoor position and walking speed in wide passages. In the experiment described in this paper, we evaluated the accuracy of our method only in 2.5m-wide passages. However, there are many wider passages in real environments. In such passages, we will be required to consider various new aspects, such as the spacing of reference lines. This is one of the limitations of our method. Also, we plan to propose a method for grasping people flow by using indoor position and walking speed information that are collected from some pedestrians in indoor passages.

REFERENCES

- [1] P. Viola, M. Jones, and D. Snow, "Detecting Pedestrians Using Patterns of Motion and Appearance," In *Proceedings of the IEEE International Conference on Computer Vision*, Vol.2, pp.734-741, October 2003.
- [2] N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection," In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, Vol.1, pp.886-893, June 2005.
- [3] A. Fod, A. Howard, and M. J. Mataric, "Laser-Based People Tracking," In *IEEE Conference on Robotics and Automation*, pp.3024-3029, May 2002.
- [4] H. Zhao and R. Shibasaki, "A Novel System for Tracking Pedestrians Using Multiple Single-Row Laser-Range Scanners," In *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, Vol.35, No.2, pp.283-291 February 2005.
- [5] J. Burke, D. Estrin, M. Hansen, A. Parker, N. Ramanathan, S.Reddy, and M. B. Srivastava, "Participatory Sensing," In *Proceedings of the 2016 Workshop on World-Sensor-Web Held at ACM SenSys'06*, pp.1-6, October 2006.
- [6] P. G. Kannan, S. P. Venkatagiri, M. C. Chan, A. L. Ananda, and L.-S. Peh, "Low Cost Counting Using Audio Tones," In *Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems*, pp.155-168, November 2012.
- [7] T. Nishimura, T. Higuchi, H. Yamaguchi, and T. Higashino, "Detecting Smoothness of Pedestrian Flows by Participatory Sensing with Mobile Phones," In *Proceedings of the 2014 ACM International Symposium on Wearable Computers*, pp.15-18, September 2014.
- [8] M. Wirz, T. Franke, D. Roggen, E. M. Kelly, P. Lukowicz, and G. Troster, "Probing Crowd Density Through Smartphones in City-scale Mass Gatherings," In *EPJ Data Science 2013*, pp.1-24, June 2013.
- [9] M. Matsubayashi and Y. Shiraishi, "A Method for Estimating Walking Speed by Using Magnetic Signature to Grasp People Flow in Indoor Passages," In *MOBIQUITOUS 2016 Adjunct Proceedings of the 13th International Conference on Mobile and Ubiquitous Systems: Computing Networking and Services*, pp.94-99, November 2016.
- [10] S. Bell, W. R. Jung, and V. Krishnakumar, "WiFi-based Enhanced Positioning Systems: Accuracy Through Mapping, Calibration, and Classification," In *Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Indoor Spatial Awareness*, pp.3-9, November 2010.
- [11] K. Kaji and N. Kawaguchi, "Design and Implementation of WiFi Indoor Localization based on Gaussian Mixture Model and Particle Filter," In *2012 International Conference on Indoor Positioning and Indoor Navigation*, pp.1-9, November 2012.
- [12] S. Sen, J. Lee, K.-H. Kim, and P. Congdon, "Avoiding Multipath to Revive Inbuilding WiFi Localization," In *Proceeding of the 11th Annual International Conference on Mobile Systems, Applications and Services*, pp.249-262, June 2013.
- [13] J. Chung, M. Donahoe, C. Schmandt, I.-J. Kim, P. Razavai, and M. Wiseman, "Indoor Location Sensing Using Geo-Magnetism," In *Proceedings of the 9th International Conference on Mobile Systems, Applications, and Services*, pp.141-154, June 2011.
- [14] K. P. Subbu, B. Gozick, and R. Dantu, "LocateMe: Magnetic-Fields-Based Indoor Localization Using Smartphones," In *ACM Transactions on Intelligent Systems and Technology*, Vol.4, No.4, Article 73, pp. 1-27, September 2013.
- [15] R. Ban, K. Kaji, K. Hiroi, and N. Kawaguchi, "Indoor Positioning Method Integrating Pedestrian Dead Reckoning with Magnetic Field and WiFi Fingerprints," In *2015 8th International Conference on Mobile Computing and Ubiquitous Networking*, pp.169-174, January 2015.
- [16] M. Kourogi and T. Kurata, "Personal Positioning based on Walking Locomotion Analysis with Self-Contained Sensors and a Wearable Camera," In *15th IEEE International Symposium on Mixed and Augmented Reality 2003*, pp.103-112, October 2003.
- [17] H. Vathsangam, A. Emken, D. Sprujt-Metz, and G. S. Sukhatme, "Toward Free-Living Walking Speed Estimation Using Gaussian Process-based Regression with On-Body Accelerometers and Gyroscopes," In *Proceedings of the 2010 4th International Conference on Pervasive Computing Technologies for Healthcare*, pp.1-8, March 2010.
- [18] J. Park, M. Patel, D. Curtis, S. Teller, and J. Ledlie, "Online Pose Classification and Walking Speed Estimation Using Handheld Devices," In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, pp.113-122, September 2012.
- [19] Y. Sakurai, C. Faloutsos, and M. Yamamuro, "Stream Monitoring under the Time Warping Distance," *2007 IEEE 23rd International Conference on Data Engineering*, pp.1046-1055, January 2007.

A proposal of method to correct GPS positioning error by using acceleration sensor

Chiaki Yashiro^{*}, Tomoyuki Yashiro^{**}

Graduate School of Information Science, Chiba Institute of Technology, Japan

^{*}chiaki.yashiro020@gmail.com, ^{**}yashiro@net.it-chiba.ac.jp

Abstract - The location acquisition is one of the major function of the smartphones to provide various services. By acquiring the detailed position, it is possible to provide a pinpoint information on where the obstacles are located and on which sidewalk has steps/stairs, and the services that can help aged, blind or disabled people to select safe and comfortable way is able to provide. However, GPS requires line of sight between terminal and satellites, large positioning error may occur in an urban area. Therefore, we examined a method to correct GPS positioning error by using acceleration sensor in order to realize the service that requires high accuracy. We focus on jerk that is not affected by the gravitational acceleration. The two type of jerk can be calculated. One is from the acceleration sensor, and another is from the GPS. By comparing two jerk degrees, a GPS position with large error can be detected. The error of the acceleration sensor is accumulated and the accuracy is relatively low. On the other hand, GPS is greatly affected by the surrounding environment, but in most areas, the accuracy is high. In the GPS high accuracy area, the accumulated error of the acceleration sensor can be corrected by GPS and GPS low accuracy area, acceleration sensor can be used to acquire the position. We combine the advantages of the both system to realize various services.

Keywords: GPS, Acceleration sensor, Jerk

1 INTRODUCTION

Smartphones are rapidly spreading. A lot of users are various services. One of those is location acquisition services. The location acquisition is one of the major function of the smartphones to provide various services. Satellite positioning service is a mainstream of location acquisition method. By receiving radio waves transmitted by many satellites, the terminal determines the position of itself and that is usually the position of the user. However, current satellite positioning services mainly uses GPS(Global Positioning System) satellites operated by USA. In the case of insufficient number of satellites can utilize to achieve the location, we can't get the accurate position. As a result, current GPS positioning system can get only the rough position including errors up to 10m in typical conditions. In the case of car navigation system, rough positioning does not cause a serious problem because the car position is basically limited on the road. However, pedestrians and bicycles use various ways on the road. They should passage left or right sidewalk, roadways, crosswalk and so on and thus the pedestrian navigation system requires detailed position. Also, the navigation system with detailed position has a great benefit to disabled people. By acquiring the

detailed position, it is possible to provide a pinpoint information on where the obstacles are located and on which sidewalk has steps/stairs, and the services that can help aged, blind or disabled people to select safe and comfortable way is able to provide. We will examine a method to correct GPS positioning in order to realize services that can't be realized with current GPS positioning accuracy. In this research, a proposal of method to detection GPS positioning error by using acceleration sensor built in smartphone.

2 RELATED WORK

In the literature [1], in order to realize pedestrian navigation that can be used indoors and outdoors, we proposed a personal positioning system integrating self-contained sensor based on walking motion, GPS, active RFID tag system. Dead-reckoning tends to lower the estimation precision of the moving direction due to the drift error of the gyro sensor and the like. In this document, the moving direction of dead-reckoning is corrected by using GPS positioning result and Kalman filter. By doing this, the influence of the drift error of the gyro sensor is eliminated. As a result, the mean absolute error with GPS alone was 16.4 m, whereas in the method of this document it is 7.2 m.

In the literature [2], the antenna radiation pattern of the smartphone is modeled by utilizing radio wave propagation analysis and electromagnetic field analysis, and the GPS positioning accuracy of the bicycle traveler is evaluated by simulation. High sensitivity receivers were greatly influenced by multipath, and showed that errors greatly changed depending on the situation.

In the literature [3], a position estimation system combining absolute position acquired by GPS and relative position by dead reckoning is proposed. The combination of both is realized by using the Kalman filter. In the method of this document, the position estimation error by GPS positioning alone is improved by 6.51 m at maximum. We showed that the accuracy of position estimation improves by correcting the position acquired by GPS with the measurement value of the sensor installed in the mobile terminal.

In the literature [4], we performed confirmation and evaluation of error characteristics by performing GPS positioning using multiple GPS dedicated devices at the same time. We collect data of 24 hours or more (Measurement result of 1400 points or more) using seven terminals. When using the same model and measuring GPS at the same place and at the same time, we assumed that the number and type of GPS satellites to be accessed are the same. However, slight differences were observed in the number of satellites supplemented and the satellites used for

error calculation. At the same time, the number of used satellites of 7 terminals and the satellite number to be used coincided with only 189 out of 1440 positions. In the survey of the variation of the error due to the variation of the number of satellites used, it was found that the positioning error decreases as the number of satellites used increases. However, it was confirmed that if the number of satellites used exceeds 6, accuracy will hardly increase even if the number of satellites increases. We also showed that the error can be reduced by adaptively selecting satellites to be used even if the number of satellites used is small.

In the literature [5], odometry and GPS were sensor-fusion using a position estimation algorithm that expresses the probability distribution of the position of the robot called MCL as particles, and the self-position of the robot was estimated. As a result, sensor fusion and self-position correction using MCL were performed sufficiently, and the effectiveness of the proposed method was shown.

In the literature [6], for the purpose of improving the accuracy of walking route estimation, we propose reduction of error and accumulation error of gyro sensor using Kalman filter and map matching. Movement path estimation was performed by three methods: no correction, Kalman filter application, and Kalman filter and map matching. The average errors are 33.0 m, 13.1 m and 0.5 m, respectively, and the effectiveness of the proposed method is shown.

In the literature [7], we propose a method to reduce multipath error using signal strength. The error of the DGPS positioning without correction was about 6 m, but the error was 4 m in the corrected DGPS positioning. Also, the above result is obtained by correcting the pseudo distance of only two satellites. It is stated that errors can be further reduced by correcting all satellites.

In the literature [8], a method for improving the positioning accuracy of neighboring receivers is proposed by using a method called mobile DGPS, in which a GPS receiver which can independently measure with high accuracy and high accuracy is a pseudo reference station of DGPS. In the case of not using the proposed method, there was an error of more than 10 m even with positioning by 8 satellites, but the proposed method showed that it can be suppressed within 2 m of error by positioning with 6 satellites.

In the literature [9], log data traveled by car using smartphone acceleration sensor, geomagnetic sensor, GPS was collected for a total of 22.3 km. Lane estimation is performed by detecting right / left turns from the log data. In the proposed method, the average correct answer rate for lane estimation was 87.1%, and the effectiveness of the proposed method could be confirmed.

In the literature [10], attempts are made to detect obstacles by measuring the traveling route using the celestial satellite system. As a result, although improvement in positioning accuracy could be confirmed by using the celestial satellite system, it did not reach the detection of obstacles. It is stated that it is necessary to further improve the accuracy of GPS.

In the literature [11], a position estimation method considering the average bias of the residual distance measurement error not considered in the conventional

position estimation in single positioning using GPS is proposed. The residual distance measurement error is thought to be greatly affected by the elevation angle of the satellite. In this paper, we propose a method to predict the residual distance measurement error by obtaining the elevation angle of the satellite, and showed the possibility of improving the positioning accuracy.

In the literature [12], we are conducting a basic study of a system to use as a pseudo satellite of road signs GNSS so that highly accurate vehicle positioning can be performed even in building streets where supplemental possible satellites are few when positioning by GNSS. As a result of the positioning experiment near the skyscraper, the positioning error due to the four satellites was 20.74 m at the minimum, but the positioning with the three satellites and the marker has an error of 4.61 m. Indicating that the sign can be a pseudo satellites.

3 PROPOSED METHOD

In this paper, we used acceleration sensor built in smartphone, we propose a detection method of large positioning error in GPS positioning. Acceleration is greatly affected by gravitational acceleration. Therefore, accurate comparison can not be made unless gravitational acceleration is properly removed. We focus on jerk that is not affected by the gravitational acceleration. The jerk is a value expressing the rate of change of acceleration per unit time, it is not affected by gravitational acceleration. The two type of jerk can be calculated. One is from the acceleration sensor, and another is from the GPS.

By acquiring the value with the acceleration sensor, acceleration in three directions of x axis, y axis, and z axis can be obtained. We combined the obtained values using the equation (1). Let x be the acceleration of the X axis, y be the acceleration of the Y axis and z be the acceleration of the Z axis. Then let a be the synthesized acceleration.

$$a = \sqrt{x^2 + y^2 + z^2} \quad (1)$$

The jerk can be calculated by differentiating the acceleration by the unit time. Calculate jerk using equation (2). The jerk is j and the time is t.

$$j = \frac{da}{dt} \quad (2)$$

The GPS can obtain the acceleration from the difference of the position information. The jerk can be obtained by using the equation (2) for the obtained acceleration. By comparing two jerk degrees, a GPS position with large error can be detected. The amount of change in jerk is difficult to compare if the peaks above and below the jerk value do not appear at the same timing on the GPS and the acceleration sensor. In order to make the change amount easy to compare, draw the envelope that connects the peaks at the top and bottom of the value. Compare the width of the envelope of each data and assume that the GPS measurement position contains a large error if the GPS envelope width is wider than the envelope of the acceleration sensor. When the jerk

moves at a constant speed, the change in value is small. However, when the jerk moves at a sudden acceleration / deceleration, the change in value is large. If the jerk calculated from the acceleration sensor changes greatly, it can be considered that the GPS measurement position has moved larger or smaller than the original advanced distance. It is considered that the GPS measurement position contains a large error. By comparing two jerk degrees, a GPS position with large error can be detected.

4 EXPERIMENTS

4.1 EXPERIMENTS ENVIRONMENT

Table 1 is overview of data collection environment. By using Zenfone2 by ASUS company to data collection. The experiment was carried out at the outer circumference of Chiba Institute of Technology Tsudanuma campus, which is about 1.5 km in distance one round at a time. We got around once per data collection. Experimental environment is shown in the Figure 1 Experimenter run constant speed at bicycle. Terminal holding position is breast pocket. Data collection application is use HASC Logger. GPS was acquired 1Hz and acceleration sensor was acquired at 100Hz. Positioning by GPS was started after waiting 30 seconds from the start of measurement because the start of measurement is unstable. Also, after the run was completed wait 10 seconds and finish the measurement. A total of about 20 hours of data was collected. As a result, a data collection was average 394.0 seconds. The collection data transfer PC. For GPS data, the jerk was calculated from the distance between measurement position. After differentiating the data of the acceleration sensor to calculate the jerk, the average is calculated every 100 jerk data and it is adjusted to the number of GPS data. After that, an envelope was drawn from the jerk of each data. These processes were performed by MATLAB which is numerical analysis software developed by MathWorks.

Table 1: Data collection environment.

Using terminal	Zenfone2 by ASUS
Terminal OS	Android5.0
Place	Chiba Institute of Technology Tsudanuma campus , Circumference
Distance	About 1.5km
Collection data	When driving a bicycle Acceleration(100Hz) GPS(1Hz)
Terminal holding position	Breast pocket
Using application	HASC Logger



Figure 1: Data collection location.

4.2 EVALUTION METHOD

We compare the width of the envelope of each data using the proposed method. Consider what percentage of the total measurement positions the error can be detected.

4.3 EXPERIMENTS RESULT

An example the result of the experimental are shown in Figure 2 and Figure 3 The horizontal axis of the graph in the figure represents time and the vertical axis represents the envelope value of jerk. The values of the change amount due to the envelope curve were compared. Errors could be detected at 202.3 positions out of 394.0 average data of all data, and errors could be detected in 51.2% of the total. Also, the maximum value of the detection ratio was 70.1%, and the lowest value was 31.9%. Figure 3 shows the measurement position detected by the proposed method. The blue dot indicates the measurement position by the GPS, and the red dot indicates the measurement position where the error is detected by the proposed method out of the GPS measurement position. The blue line seen in part A is a straight line connecting GPS measurement positions. From this, it can be confirmed that the error can be detected at the measurement position greatly deviated at the start of measurement like the portion A and the measurement position deviated from the road. However, there are also measurement positions that are not detected as despite being greatly out of the way as in section B. The causes of such detection omissions are shown in the figure. The proposed method detects when the jerk calculated from the distance between two GPS points does not indicate constant speed movement. Therefore, as shown in the left figure, when the measured position appears as a single unit, the jerk indicates acceleration / deceleration and it can be detected. However, as shown in the figure on the right, when the measurement position is entirely out of position, no abnormality appears in the distance between the two points. Even if the jerk indicates a constant speed movement, it can't be detected even if it is out of the way. As a result, we lost detection as shown in part B of the figure.

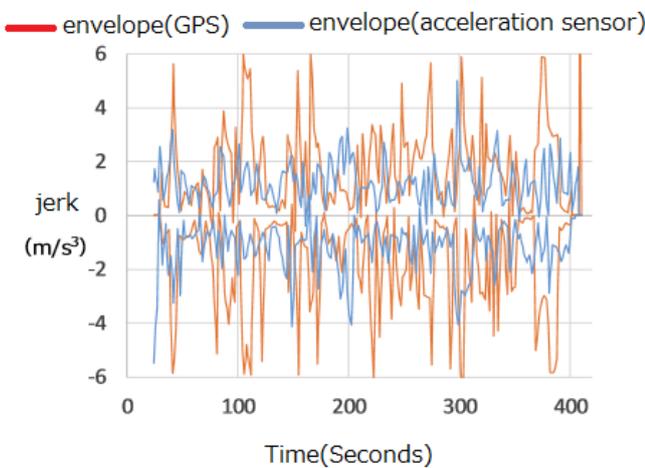


Figure 2: Comparison of envelopes of jerk.



Figure 3: Proposed method detection position.

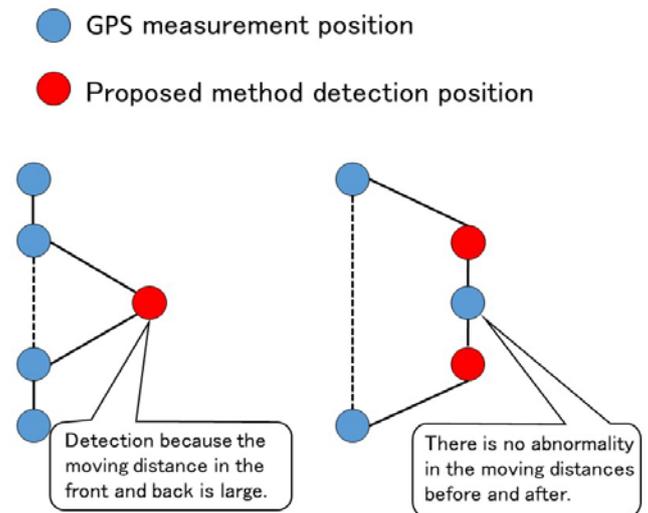


Figure 4: Cause of detection omission.

5 CONCLUSION

The penetration rate of smartphones is increasing year by year, and services using location information are also frequently used. In recent years, position information is utilized not only in navigation but also in various services such as games and photographs. As a means for acquiring the position of the terminal, a satellite positioning system such as GPS is used. However, in the present situation, there are cases where the error factor is large and the position including a large error is sometimes measured. Therefore, we believe it is difficult to provide services using detailed location information with the current positioning accuracy as it is. Therefore, we examined a method to correct GPS error. In this research, we have investigated the detection method of GPS measurement position including large error by using the acceleration sensor built in many smart phones. By comparing the jerk calculated from the distance between the two points of the GPS measurement position and the jerk calculated from the value of the acceleration sensor, a measurement position with a large error was detected. This is because the jerk can be compared without considering the gravitational acceleration which greatly influences the value of the acceleration sensor. When the jerk is moving at a substantially constant speed, the change in value is small, but the value changes greatly with rapid acceleration / deceleration. If the jerk calculated from the GPS changes greatly despite the fact that the jerk calculated from the acceleration sensor indicates the constant speed movement, the GPS measurement position seems to contain a large error. Therefore, we considered that GPS measurement positions with large errors can be detected by comparing jerk of each data. We compare jerk by the value of the envelope width.

As a result of comparing the amount of change in jerk of data respectively, it was possible to detect the error at the position of 51.2% of all measurement positions. From this fact, it was confirmed that there was a possibility of correcting about half of the measurement positions out of all measurement positions. Since this proposal calculates the jerk of GPS from the distance between two points of GPS, it is good for detecting abnormal value appearing alone. However, it was also confirmed that it was difficult to detect when the measurement position was deviated as a whole.

When correcting the GPS using the actual acceleration sensor, it is considered that more effective correction can be performed by adjusting the weights of the values of the GPS and the acceleration sensor using this proposal. For that reason, it is necessary to confirm the validity of this proposal by confirming the detection omission and detection of this proposal from now on. In addition, it is necessary to reconsider the detection method so that it can be detected in various situations.

REFERENCES

[1] M. Kouroggi, N. Sakata, T. Okuma, and T. Kurata, Indoor/outdoor pedestrian navigation with an embedded GPS/Rfid/self-contained sensor system, ICAT'06 Proceedings of the 16th international

- conference on Advances in Artificial Reality and Tele-Existence, pp.1310-1321 (2006).
- [2] R. Furukawa, T. Matsumoto, M. Asanuma, Y. Emori, A. Sallehuddin, Y. Kishiki, Evaluation of Accuracy of GNSS Positioning for Bicycles in Urban Canyon Situations Using Electromagnetic and Radio Propagation Simulation, IEICE, Vol.116, No.216, ITS2016-13, pp.39-43 (2016). (in Japanese)
- [3] R. Ogawara, M. Fujii, H. Hatano, and Y. Watanabe, A study on location tracking system using Kalman filter based on sensor information, 2012 International Symposium on Information Theory and its Applications, pp.184-188 (2012).
- [4] Y. Odaka, S. Takano, Y. In, M. Higuchi, H. Murakami, The evaluation of the error characteristics of multiple GPS terminals, CSCS '11 Proceedings of the 2nd international conference on Circuits, systems, control, signals, pp.13-21 (2011).
- [5] T. Ikeda, F. Nakazawa, Y. Ando, and M. Mizukawa, Research on Sensor Fusion and Position Estimation of Autonomous Robot in Outdoor Environment using Particle Filter, Robomec 2009, 2A2-E20(1) - 2A2-E20(4) (2009). (in Japanese)
- [6] S. Tagawa, I. Takumi, and T. Uchiya, Study on improvement of pedestrian path estimation accuracy using Kalman filter and map matching, Proceedings of the 75th National Convention of IPSJ 2013(1), pp.207-208 (2013). (in Japanese)
- [7] S. Takashi, N. Kubo, and A. Yasuda, The possibility of the precise positioning and multipath error mitigation in the real-time, The 2004 International Symposium on GNSS/GPS (2004).
- [8] T. Kitani, and H. Hatano, A cooperative positioning method to improve the accuracy of neighboring GPS receivers, IPSJ SIG Technical Report, Vol.2012-ITS-49, No.2 (2012). (in Japanese)
- [9] T. Makino, Y. Ito, and Y. Shiraishi, A Method of Traffic Lane Estimation with a Smartphone, IPSJ Journal, Vol.55, No.2, pp.812-825 (2014). (in Japanese)
- [10] Y. Izumi, PROBER proposal to dynamically collect information using smartphone, Chiba Institute of Technology, Graduation thesis (2015). (in Japanese)
- [11] T. Hattori, M. Fujii, H. Hatano, A. Ito, and Y. Watanabe, A study on residual ranging error for GPS positioning, IEICE, Vol.116, No.216, ITS2016-12, pp.33-38 (2016). (in Japanese)
- [12] R. Takushi, and T. Nagaosa, A Study of Vehicular Localization Using Traffic Sign Information and GNSS, IEICE, Vol.115, No.504, ITS2015-93, pp.53-58 (2016). (in Japanese)

A BLE Beacon's Movement and Equipment Defect Estimation Method Based on Comparison of Room-level Wi-Fi and BLE Fingerprints

Shota Ikeda[†], Fumitaka Naruse[‡], and Katsuhiko Kaji[‡]

[†]Graduate School of Business Administration and Computer Science, Aichi Institute of Technology, Japan

[‡]Department of Information Science, Aichi Institute of Technology, Japan
{b17703bb, k14092kk, kaji}@aitech.ac.jp

Abstract - In this research, we use the room Wi-Fi model and BLE model to detect problems with BLE beacons such as movement, battery outage or other malfunctions. When using a beacon for living space information, it is necessary to install one or more in each room. It is considered difficult for an administrator to manage all of the beacons installed in each room at the same time. In the proposed method, a room model including a Wi-Fi model and a BLE model is created. When entering a room, we observe the Wi-Fi and BLE information in that room, and compare each model to detect beacon defects. We conducted an experiment to test the accuracy of room estimation and beacon defect estimation. For isolated rooms, estimation accuracy was 90% or greater. defect such as battery exhaustion, as well as movement of the beacon, could also be detected from the room estimation result. When the beacon observable but are not with high accuracy of 94%. However, when the beacon should not be observable, accuracy was as low as 40%.

Keywords: BLE beacon, fingerprint, room estimation, defect estimation, Wi-Fi

1 Introduction

BLE beacons (from here on, simply beacons) are used for various applications such as indoor positioning, room recognition, and stamp-rally applications of public facilities [1][2].

When using beacons for room information, it is necessary to install one or more in each room. It is considered difficult for an administrator to manage all of the beacons installed in each room at the same time. For this reason, it is necessary to inspect the beacons installed in each room one by one, but such inspection is costly. Also, when a beacon has been moved, it is considered difficult find the new location.

We propose a beacon monitoring method. Possible beacon defect is not only battery defect. Most beacons are small and hard to fix, so that someone could easily move one to another room or take it out entirely. Our method can find such kinds of beacon defect even when the system is in operation. To achieve monitoring, we adopt crowd-sensing technology. There are multiple users of the beacon-based room estimation application. When detecting a room, received signal information is gathered. Additionally, we utilize Wi-Fi and BLE fingerprints simultaneously to find beacon defect.

We detect defects such as battery exhaustion, defect, and movement as possible malfunctions of the beacon. Since the beacon is battery-operated, the battery runs out during long-term operation. Also, beacons are small in size and difficult to

fix, so there is a possibility that they are moved accidentally. In addition, there are cases where beacons are mistakenly installed in the wrong rooms.

Using the proposed method, it is possible to automatically check the operation of a beacon. It is not necessary to periodically check the operation of each beacon. In addition, since the beacons are managed all at once, it is possible to operate without relying on a person in charge of each room.

The rest of this paper is structured as follows. In Section 2, we explain related research. Section 3 gives an overview of the proposed method. Section 4 describes experiments using the proposed method. Section 5 summarizes and discusses future issues.

2 Related research

Wireless LAN, beacons, built-in sensors of smart phones, etc. are often used for indoor position estimation and room estimation research. However, there are few studies on detecting location or on terminal management.

There is research on the behavior monitoring of BLE beacons using participatory sensing. In Asahi's method [2], the beacon is used as a check point of the electronic stamp rally. When the smartphone terminal that introduced the application receives radio waves, beacon information and time data are sent to the server. When monitoring the information transmitted to the server, the data continuously transmitted may be interrupted in some cases. Since the transmission of data suddenly stops, it is a method that grasps the activity state of the beacon. This method is considered to be effective as beacon management during service operation. However, if the beacon installed as a checkpoint is moved, or if it is installed at a place where people do not always go, it cannot be determined whether the beacon is running or not.

In this method, if someone brings a smartphone within range of a beacon, it can be confirmed whether it is in operation.

When estimating position using campus LAN, it is necessary to measure the radio wave intensity of Wi-Fi for each position. When using the wireless access point to estimate the position as in Dhruv's method [3], it is necessary to determine the measurement position of the radio wave intensity considering the base station. In contrast, our method creates Wi-Fi and BLE fingerprints. Instead of a fingerprint for each location, it creates a fingerprint for each room based on the measured data. When you create a fingerprint for each room, you cannot figure out where you are in the room. However, it becomes easy to grasp whether or not you are in that.

3 Proposal of beacon management method using radio condition of Wi-Fi and BLE

In this research, we use the Wi-Fi model and BLE model of a room as a proposed method, and detect defects such as battery outage or breakdown of a beacon and movement of a beacon. An outline of this method is shown in Fig. 1. In scene 1, when comparing the observed data with the room model of room α , BLE signal is not received from beacon A. Therefore, beacon A is thought to have experienced defect such as battery exhaustion or defect. In scene 2, when comparing the model of room α with the observed data, the existence of beacon B can be confirmed in the observed data. Therefore, since there is a high possibility that beacon B was moved from room β to room α .

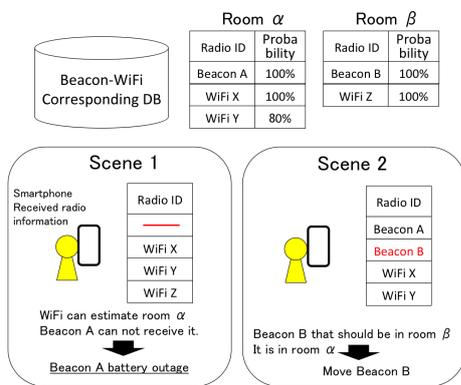


Figure 1: Outline of proposed method

The reason for combining Wi-Fi and BLE is to use Wi-Fi to detect a beacon defect when a problem occurs in the beacon. It also plays the role of room estimation. When estimating a room by installing a single beacon in a small room such as a small classroom or a laboratory, it is impossible to estimate a room if a beacon fails or a beacon is taken out. Also, in case of defect, it will be repaired or replaced. However, when it is brought out, it is necessary to search for a beacon terminal.

During operation, we observe data from various people. Basically, the room estimation is performed using BLE data, but room estimation is performed simultaneously with BLE using Wi-Fi data at regular intervals. This is to check the status of the beacon on a regular basis. when room estimation is performed normally, each model is updated using the observed Wi-Fi and BLE data.

3.1 Data collection and assumption

In the proposed method, the fingerprint for Wi-Fi and BLE is collected in advance. As a premise, smartphones are used to collect data. Observation data at the time of preliminary collection are gathered with correct answer room name known. In data collection, we walk around in each room, and record radio observation information at 10 second intervals.

In a general Wi-Fi fingerprint collecting method, radio waves obtained while stationary for several seconds are regarded as the fingerprint at that position[4][5]. On the other hand, in the proposed method, we model the radio wave environment of

each room. Therefore, electric wave information of the whole room is necessary. Therefore, we walk around the room and collect fingerprints in various places in the room.

Data collection is conducted by participatory sensing, and when a user with a smartphone enters a room, observation is assumed simultaneously with room estimation. Basically, the smartphone collects BLE data, but also collects Wi-Fi data at regular intervals. We integrate the collected Wi-Fi of the room and data of each BLE, create a Wi-Fi model and BLE model for each room, and make a room model.

3.2 Modeling observation data

We create the first model as a reference for room estimation. Next, we create Wi-Fi and BLE models for each room based on the data observed as in Figure 2. In the proposed method, for the radio waves observed in the room, the probability of observing the radio waves is obtained, and the Wi-Fi model and BLE model are generated.

Also, for both the Wi-Fi model and BLE model, radio waves with observation probability of less than 50% are not included in the model. Room estimation is performed based on the observation probability of each radio wave. However, if extremely low radio wave information is included at that time, the probability of room estimation is considered to be low. This is because radio wave information, which is not frequently observed, is used for room estimation.

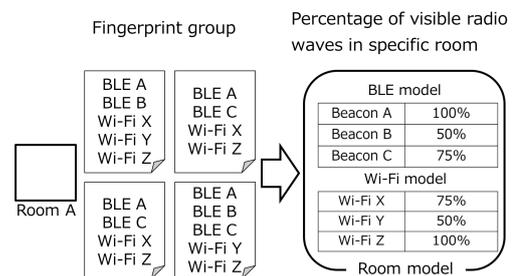


Figure 2: Model creation method

3.3 Room estimation

Room estimation during operation is explained next. We compare the data observed at a certain timing, the list of Wi-Fi and BLE in each room built in advance, the BLE model and the Wi-Fi model, and estimate the room. Here, we will explain room estimation by the BLE model, but room estimation using the BLE model and that using the Wi-Fi model are done in the same way.

Let O_b be the set of BLE radio waves contained in observational data O . At this time, the probability $p(r)$ existing in a room r is calculated as follows. Here, we denote the observed probability of radio wave a in room r as $p(a|r)$ and the set of radio waves contained in the BLE model of room r as M_b^r . First, we obtain the set of radio waves common to O_b and M_b^r as $O_b \cap M_b^r$. Also, we obtain a set of radio waves that are included in M_b^r and not included in O_b as $M_b^r - (O_b \cap M_b^r)$. Next, we obtain the probability that the radio waves of the set element can be observed in the room r as follows.

$$p(r) = \prod p(a|r) \times \prod (1 - p(b|r))$$

Where, $a \in O_b \cap M_b^r, b \in M_b^r - (O_b \cap M_b^r)$

Comparing the observation data with any room model as described above, the probability of being a specific room is required. Let the room with the highest probability be the room where the smartphone is currently.

3.4 Beacon defect estimation

Based on the results of BLE room estimation and Wi-Fi room estimation, we compare BLE radio wave O_b with the BLE model and estimate the defect of a BLE beacon. The malfunction of a BLE beacon is such that radio waves are not transmitted due to battery exhaustion or breakdown, or it has been taken out of the room. For these defects, we will discover two types of inconsistencies: Should not be observable; and patterns that are observable but are not. Then, we perform defect analysis.

As a precondition, R_b is the room estimated by the BLE model. R_w is the room estimated by the Wi-Fi model. O_b is the set of BLE beacons received at a given observation. M_b^R is the set of BLE beacons included in the BLE model in room R.

First, we will show the algorithm for finding the beacon set E_{mh} (mh means move here), which is supposed to be unobservable. What can be found with this pattern is that the beacon was moved to a room observed from some room.

Suppose the room estimate R_b based on the BLE model is correct. The beacon set that is supposed to be impossible to observe can be obtained as follows.

$$E_b^{mh} = O_b - (M_b^{R_b} \cap O_b)$$

On the other hand, suppose that the room estimate R_w based on the Wi-Fi model is correct. The beacon set that is supposed to be impossible to observe can be obtained as follows.

$$E_w^{mh} = O_b - (M_b^{R_w} \cap O_b)$$

Here, if R_b and R_w are different, elements of E_b^{mh} and E_w^{mh} are also different. In that case, their union is regarded as a candidate for a problem.

$$E^{mh} = E_b^{mh} \cup E_w^{mh}$$

Next, we show an algorithm to examine the beacon set E^{mt} (mt means move to somewhere), which is supposed to be observed but is not observed. What we can discover with this pattern is a malfunction, a battery exhaustion, or that the beacon has been moved out of the observed room.

Suppose the room estimate R_b based on the BLE model is correct. A beacon set that is supposed to be observed but is not observed is obtained as follows.

$$E_b^{mt} = M_b - (M_b^{R_b} \cap O_b)$$

On the other hand, suppose that the room estimate R_w based on the Wi-Fi model is correct. A beacon set that is supposed to be observed but is not observed is obtained as follows.

$$E_w^{mt} = M_b - (M_b^{R_w} \cap O_b)$$

Here, if R_b and R_w are different, elements of E_b^{mt} and E_w^{mt} are also different. In that case, their union is regarded as a candidate for a problem.

$$E^{mt} = E_b^{mt} \cup E_w^{mt}$$

4 Room estimation and beacon defect estimation experiment

By examining the room estimation method conducted in Section 3, we can verify the accuracy of room estimation. We also conducted experiments as to whether beacon defects could be detected. The factors considered to affect the detection accuracy may be the case where the size of the room, the position of the room, the number of beacons, and multiple defects occur simultaneously.

As the first experiment, after creating the room model, we observe the Wi-Fi and BLE data in each room and obtain the accuracy of room estimation. Since beacon radio waves are weaker than Wi-Fi, it can be considered observable in the room and still not be observed. Therefore, we observe data at various places in the room. We believe the probability of room estimation will change with the distance of the room. In addition, it is thought that the probability of room estimation will be low if all the beacons with low radio field intensity are used.

As a second experiment, we will detect malfunctions assuming beacon movement, defect etc. We anticipate the following kinds of defect: one cannot observe the beacon in the room; the beacon in the room has malfunctioned; a beacon that should not be observable originally is detected; and a beacon installed in one room has been moved to another room.

4.1 Experiment setting

As the experimental setting, data collection is done in each room using a smartphone as in Section 3. The room used in the experiment is shown in Figure 3. Wi-Fi access point is installed in each room. Measurement of Wi-Fi and BLE data is performed once every 10 seconds, 10 times in total. Data collection takes place everywhere in the room.

Everywhere in the room at least one beacon is signal can be received. All UUIDs are unified. In room A, three beacons are installed, and it is apart from the other rooms. In room B, two beacons are installed, and room B is in a different building from the other rooms. In room C, two beacons are installed, and it is next to room D. In room D, one beacon is installed in the center. Room E is a large lecture room and has beacons installed three places.

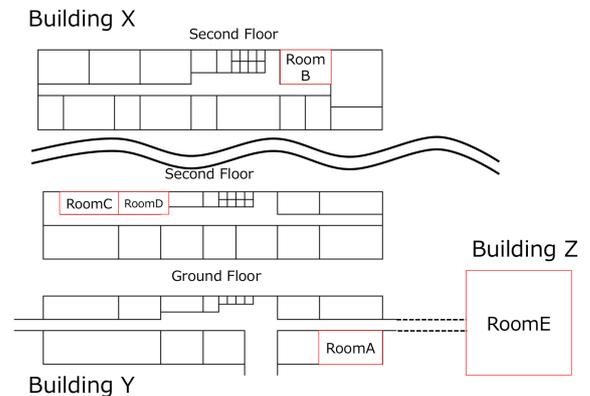


Figure 3: Floor map

4.2 Room estimation experiment

In the room estimation experiment, we used the Wi-Fi and BLE observation data and experimented on how accurately room estimation could be estimated. If the accuracy of room estimation is low, it is thought that estimation of beacon defects will be affected. Since there is a possibility that the experiments to be performed next may be affected, the boundary value is also examined. Boundary values of -50 dBm, -75 dBm, and -90 dBm were used for room estimation. Experimental results are shown in Table 1.

From the experimental results, we could estimate a room with high probability except for ones with a room next door. Moreover, accuracy is high when room estimation is performed at -75 dBm from the experimental result. Therefore, the boundary value of the defect estimate is -75 dBm.

As a result, the influence of the number of beacons and the size of the room is small. However, when the position of the room is close, the beacon radio waves can be acquired, so the detection accuracy is low.

Table 1: Room estimation result

Room estimation	-50dBm		-75dBm		-90dBm	
	BLE	Wi-Fi	BLE	Wi-Fi	BLE	Wi-Fi
Room A	100%	100%	100%	100%	100%	100%
Room B	100%	100%	100%	100%	100%	100%
Room C	100%	70%	100%	80%	100%	100%
Room D	30%	100%	40%	100%	30%	80%
Room E	100%	100%	100%	100%	100%	100%
Overall probability	86%	94%	88%	96%	86%	96%

4.3 Beacon defect estimation experiment

In the defect estimation experiment, a beacon defect was detected, assuming a malfunction such as battery exhaustion, defect or movement of a beacon. The verified defect situation is shown as follow. In defect 1, one of three beacons in room A is supposed to be in a situation where it cannot be used due to battery outage or malfunction. In defect 2, the beacons of room B and room C were arranged by mistake, or moved. In defect 3, it is assumed that the beacon of room E has been moved to room D. Experimental results are shown in Table 2.

Beacons can be observed. However, the accuracy that the beacon can not be observed was 94%. The beacon should not be observable, but the estimation result was about 40% of the correct answer of the defect estimate.

As a consideration, is these results, if the room is isolated and not affected by other beacons, defect estimation can be done with high probability. However, it was confirmed that the probability of estimation of defects is about half in the case of adjacent rooms and rooms influenced by beacons close by.

5 Conclusion

In this paper, we proposed a method of fault and relocation estimation of BLE beacons based on comparison between WiFi and BLE fingerprints for each room. We modeled the observed Wi-Fi and BLE data. Basically, we compared the BLE model with the observed BLE list and estimated the

Table 2: Defect experiment result

	Should not be observable	observable but are not
Number of observations	50	50
Total number of defect estimations	106	51
Number of correct answers for defect estimation	43	48
Incorrect number of defects detected	63	3
Correct answer rate	40%	94%

room. We also compared Wi-Fi models and Wi-Fi lists at regular intervals.

A room estimation experiment was conducted. The probability of room estimate succeeding is 91% on average. In addition, the boundary value of the radio field intensity that does not affect beacon defect estimation was -75 dBm.

We conducted a beacon fault estimation experiment, and it was possible to detect a beacon in which a malfunction had occurred. However, it was difficult to estimate defect in a room with a neighboring room or a room with a small space. The reason for this is that beacons are confused because they are more vulnerable to nearby BLE radio waves.

As a future work, a room distribution estimation will be performed by using a normal distribution of the radio wave intensity and using the strength range of the radio wave intensity. this method calculates the probability by observation that can be observed by the base station and estimates the room. Therefore, the probability of room estimation is low. By performing a normal distribution, it is expected to understand the range of the radio wave intensity seen in a specific room, accurately estimate the room even in the adjacent room, and to improve the estimation accuracy of the beacon defect.

REFERENCES

- [1] Arzad, A.K., Suleep, K.B., Bongjhin, S. Hybrid location tracking in BLE beacon systems with in-network coordination. *Consumer Communications & Networking Conference (CCNC), 2016 13th IEEE Annual*, 2016.
- [2] Asahi, D., Yokohata, Y., Inoue, T., Maeomichi, H., Tsutsui, A. Performance evaluation of user participation type BLE beacon operation monitoring in wide area electronic stamp rally experiment. *IEICE technical report Vol.115 No.466*, 2016(In Japan).
- [3] Dhruv, P., Ravi, J., E., L. Indoor location estimation using multiple wireless technologies. *Personal, Indoor and Mobile Radio Communications, 2003. PIMRC 2003. 14th IEEE Proceedings on*, 2003.
- [4] Arsham, F., Jiwei, L., Mahesh K.M., Francisco J., G. A Microscopic Look at WiFi Fingerprinting for Indoor Mobile Phone Localization in Diverse Environments. *2013 International Conference on Indoor Positioning and Indoor Navigation*, 2013.
- [5] Mohd Nizam, H., Sukhan, L. Indoor Human Localization with Orientation using WiFi Fingerprinting. *Proceedings of the 8th International Conference on Ubiquitous Information Management and Communication Article No. 109*, 2014.

A Prototype of a Precision Forecasting System for Real-time Navigation with RTK-GNSS

Areeyapinun Tiphath[†] and Tomoya Kitani[†]

[†]Graduate School of Integrated Science and Technology, Shizuoka University, Japan
{tiphath-a, t-kitani}@kitanilab.org

Abstract - In this paper, we propose a precision forecasting system for RTK-GNSS. This precision forecast system is based on a navigation satellite constellation database and a local 3D map. RTK-GNSS can provide centimeter-level accuracy positioning. The idea of a precision area forecast system comes from the fundamental problem of RTK-GNSS navigation: it needs to receive direct signal from LOS (line of sight) satellites to calculate precious position and multipath signals from NLOS (non-line of sight) satellites make accuracy worse easily. In recent years, the GNSS research community has demonstrated the algorithms to improve positioning result. Those are working on real-time navigation or post-processing, to detect and remove NLOS satellites for helping improve its results, but they are not used to predict the navigation availability and accuracy at the navigation time. This forecast system helps calculates and delivers where is capable for high precision positioning, especially in an urban area. The proposed system will detect and remove NLOS satellites for positioning calculation, and will score the predicted accuracy with remaining LOS satellites and their alignment in the sky. To support the idea of the proposed system, we have performed an experiment in Hamamatsu city. The results significantly improve in position solution and accuracy, and the proposed system can obtain 3D obstacle map for RTK-GNSS from 3D building map enhanced by observed SNR.

Keywords: GNSS, Real-time Kinematic, Precision Forecasting.

1 INTRODUCTION

In the era of multi-Global Navigation Satellite Systems (GNSS), most of navigation systems are designed with GNSS as their base. That makes impact to an industrial development and economic in everywhere around the globe. It is expected that in 2020 will be operating for about three billion satellite navigation receivers [1]. The main idea of GNSS positioning is to measure distances between navigation satellites and user receivers. There are so many positioning techniques available, depending on the calculation and desired accuracy. Real-Time Kinematic (RTK) is a promising precious positioning technique that used broadcasted reference data from a local GNSS base station.

RTK is capable for applications that require precision accuracies. RTK positioning uses carrier-phase measurement of satellite signals instead of code-phase, and thus its positioning accuracy is much sensitive about measurement error. The base station broadcasts its carrier-phase measurement of each satellite signal to local mobile stations (rovers), and the

rovers correct their own observed carrier phase measurements with the broadcasted data [2]. The accuracy of each position computation can achieve in centimeter level. It can also be observed that the results of RTK are not stable in all situations. The accuracy of RTK navigation can be affected by following factors such as operation range, satellite elevation mask, the number of visible satellites and signal obstructions from buildings, tree, radio interference, etc.

Multipath is a satellite-emitted signal arrives at the receiver by more than one path as shown in Fig. 1 [3]. This effects mainly caused by signal reflection and diffraction from various surfaces near a receiver. In an urban environment, GNSS receivers easily receive multipath signals. These multipath effects eliminate the number of received satellites and degrade the quality of the navigation [6]. We address the problem of RTK navigation in an urban environment, where line of sight (LOS) navigation satellites are few. A good number of LOS navigation satellites and good signal strength that the receiver can be received, let precision positioning be practical with RTK-GNSS.

In this project, we present a precision forecasting system. This system helps calculates and delivers which place is capable for high precision positioning, especially in an urban area. The proposed system will detect and remove NLOS satellites for positioning calculation, and will score the predicted accuracy with remaining LOS satellites and their alignment in the sky. The system pilots in Hamamatsu city, Shizuoka, Japan.

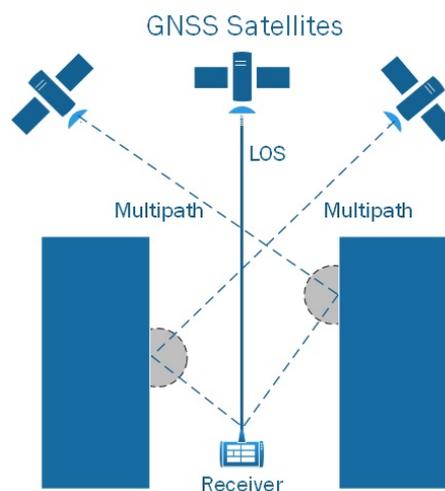


Figure 1: A multipath reflection

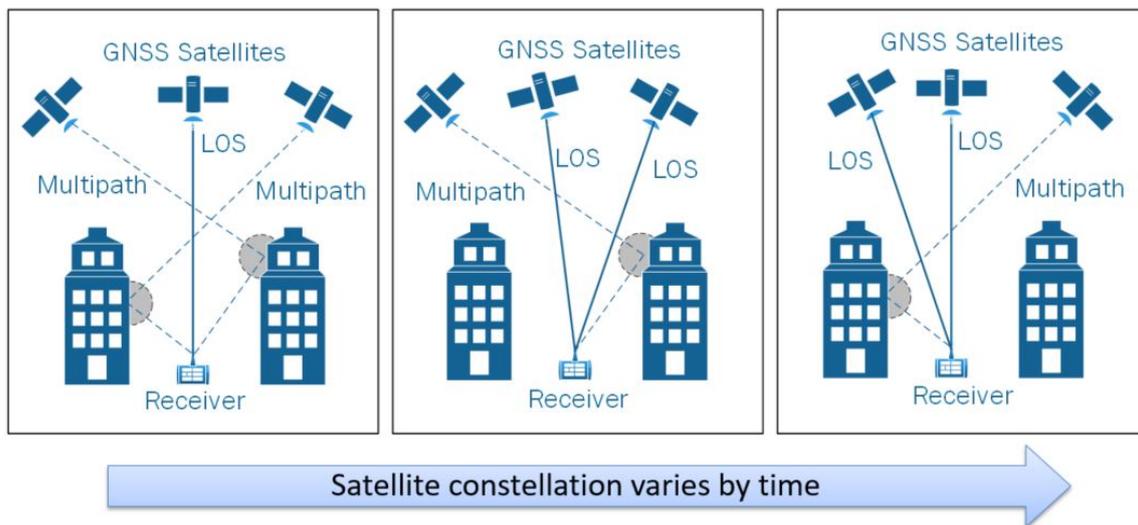


Figure 2: The visible satellites are varied by time

2 RELATIVE POSITIONING WITH GNSS AND RELATED WORK

2.1 Real-time Kinematic Method

In GNSS, a receiver can calculate its position by using an information of satellite constellations and measured distances to the satellites. The receiver needs at least four visible satellites to solve four unknown values for position determination [7]. Relative positioning is one of a positioning technique. In relative positioning, minimum of two receivers are involved in the operation. One of the receivers used as a known position during the operation, called base station. The objective of this technique is to determine the position of the other receivers, called rovers. All rovers are relative to the base station. Both receivers observe the same constellation of satellites at the same time. In other words, the two receivers record very similar errors and since the base's position is known, correction information can be generated there and distributed to the rover for solution improvement [8]. Real-time kinematic (RTK) is one of a method in the relative positioning, that use to measure the rover positioning even while it is in motion.

RTK system typically returns results in two different modes: Fixed and Float solutions. The RTK algorithm calculate the exact number of the radio wavelengths between the satellites and the base station. This process known as ambiguity resolution. In a fixed solution, the algorithm is constrained to yield a whole number in the resolution. In a float solution, the ambiguity can be a decimal or floating-point number [4]. A fixed solution typically generates precise coordinates in centimeter-level and up to 5 meters for float solution [5]. A low number of LOS satellites, poor satellite constellation geometry (GDOP) and a bad data link between the base station and the rover may prevent a fixed solution.

2.2 Related Work

In recent years, the GNSS research community has demonstrated the algorithms to improve positioning result. Example

of the algorithms are as following. NLOS satellites selection methods based on using a fisheye view images and the ratio value provided during the ambiguity resolution process [6]. Improving urban positioning accuracy using a 3D city model with optimized visibility scoring [11]. Multipath Detection with 3D Digital Maps for Robust Multi-Constellation GNSS/INS Vehicle Localization in Urban Areas [19]. Real-time GNSS positioning improvement by using crowdsourced satellite SNR measurements [17]. Evaluation of Multi-GNSSs and GPS with 3D Map Methods for Pedestrian Positioning in an Urban Canyon Environment [18].

Many algorithms using 3D mapping-aided GNSS have been developed and the results were impressive. 3D mapping-aided GNSS can vastly improve positioning accuracy in urban areas and is practical to implement [12]. A 3D map is becoming more accurate and widely available from private organizations and government. We use this data to calculate the shadows of buildings and detect LOS satellites in an urban area. In open areas, 3D map only improves the vertical position solution [12]. Those are working on real-time navigation or post-processing, for helping improve its results, but they are not used to predict the available number of LOS satellites on the navigation time.

3 PROBLEM STATEMENT AND APPROACH

During RTK operation, the following factors are the potential problems that affect RTK accuracy: the number of visible satellites, the multipath signals, the elevation mask, the ionospheric activity and the radio interference [9]. In this paper, we consider two of those factors as the main priority. The first is the number of visible satellites. During the operation, a minimum of four navigation satellites are required to estimate the receiver position and time. When RTK solution has been initialized, a minimum of four navigation satellites must be maintained to produce an RTK solution continuously. A good number of LOS navigation satellites that the receiver

can be received, lets precision positioning be practical with RTK-GNSS. The second is the multipath signals. This factor created by a signal that coming to an antenna in different directions, which result from signal reflections on various surfaces [10], especially in an urban area. Strong multipath reflections degrade the quality of an observation data and affects both reliability and availability of the solutions [6].

As mentioned in Sec. 1, an availability of RTK-GNSS navigation is limited and varies by place, time and situations. The most important factor for RTK-GNSS solution is the LOS signals from GNSS satellites. Visibility of GNSS satellites is different depending on time and place because each satellite has its own constellation and orbital period. In some situations, the visible satellites are blocked by obstacles near the receiver. Without the LOS signals from four or more visible satellites, an accurate position solution of RTK-GNSS cannot be determined [11].

It is worth nothing to make an RTK navigation at the time that LOS satellites are not available or not enough to calculate. In this paper, we propose a precision forecasting system for RTK-GNSS. The proposed system will detect and remove NLOS satellites for positioning calculation, and will score the predicted accuracy with remaining LOS satellites and their alignment in the sky.

4 A PRECISION FORECASTING SYSTEM

4.1 Concept

The idea of a precision area forecast system comes from the fundamental problem of RTK-GNSS navigation: it needs to receive direct signal from LOS (line of sight) satellites to calculate precious position and multipath signals from NLOS (non-line of sight) satellites make accuracy worse easily (Fig. 1).

There are two reasons to explain these effects. The first reason is when multiple signals are received: LOS and NLOS. Such errors can be eliminated by considering the SNR of the received signals. The second reason is when the receiver received only NLOS signal. In this case, the error magnitude might surge to several hundred meters especially in a case where surrounding structures have highly reflective surfaces [13].

In general location, the visible GNSS satellites will be different depending on time by their characteristic constellations (Fig. 2). Such a reason may limit a receiver receive the LOS satellites for precious position calculation. A situation in an urban area is worse. A lot of obstacles, such as tall buildings, block the direct LOS signals and reflect the multipath signals to the receiver. With little number of direct signals, an accurate position cannot be determined [11].

We have conducted an experiment in an urban area, to confirm that positioning result will improve by removing the NLOS satellites. We have recorded an observation data between Building No. 1 and No. 2 of Faculty of Informatics, Hamamatsu Campus for 24 hours on January 30, 2017. SNR of received signals are plotted in Fig. 3 with a fisheye camera's photo.

Then we did a post-processing on the recorded data by RTKLIB[14], a well-known open source software for RTK

calculation. From default configuration of RTKLIB, Elevation mask set to 15 degrees and SNR filters are not set. As the result of RTKPOST, we have got 3.5% of the positions as Fix solution and 96.5% as Float solution (Fig. 3 (a)). Then we changed the configurations, elevation mask from 15 to 45° and SNR from "not set" to 35 dB. Those values came from the NLOS detection result that we will discuss in Sec. 4.2. The FIX solution gained from 3.5% to 21.6% (Fig. 3 (b)). FIX solution and accuracy are significantly improved. We can conclude that NLOS satellites are effect to the solution result.

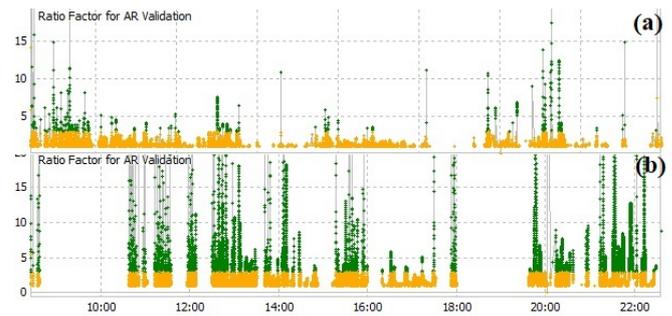


Figure 3: The green points show the results of FIX solution: (top) Default RTKLIB configuration, (bottom) Elevation mask 45 deg. and SNR 35dB

To predict a limited availability of RTK-GNSS navigation in an urban area, we propose the prototype of a precision forecasting system. As mentioned in Sec. 2.2, other algorithms do not support a forecasting of the LOS satellites. This forecast system helps calculates and delivers which place is capable for high precision positioning, especially in an urban area. The proposed system will detect and remove NLOS satellites for positioning calculation, and will score the predicted accuracy with remaining LOS satellites and their alignment in the sky.

The inputs of this system are time, position, the constellation of GNSS and 3D map of local buildings. The former two inputs are given by a user. The constellation is provided from governments that operate the GNSS. The 3D map is provided from an existing opensource data such as OpenStreetMap. Its output will be a predicted precision in a scoring scheme of the time and the place. The forecast system will search the input time and place, then map with the GNSS constellation to find available satellites. A 3D map will be used to calculate an obstacle boundary. A geometry of all available satellites will be calculated as GDOP (Geometric dilution of precision) value. The final step, the forecast system will evaluate the result in a scoring scheme and predicted the navigation.

The flowchart of the forecast system algorithm is shown in Fig. 4.

4.2 Concept Experiments

This section describes the method to find the values in each forecast system process, as list below.

- Finding of the visible satellites

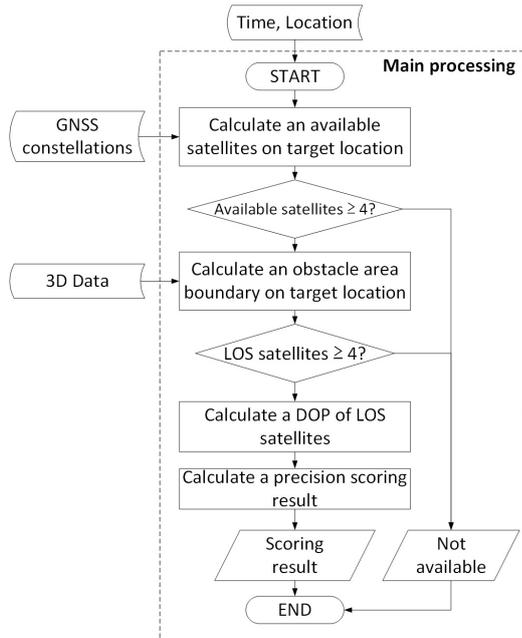


Figure 4: A flow chart of the forecast system algorithm

- NLOS satellites detection by obstacle area boundary
- GDOP calculation of LOS satellites
- Calculation of the scoring result

4.2.1 Finding of the Visible Satellites with Low-cost Logger Unit

We have implemented a low-cost prototype of a rover: a logger unit (Fig. 5). The prototype consists of a low-cost single-board computer, a low-cost RTK-ready GNSS receiver module and an antenna to record GNSS observation data directly. It can record navigation satellites signals from GPS (Global Positioning System from U.S.), QZSS (Quasi-Zenith Satellite System from Japan), Galileo (Navigation Satellite System from Europe) and BeiDou (Navigation Satellite System from China). It can also record the carrier-phase data of each of the signals.



Figure 5: A Low-cost Logger Unit

We have recorded an observation data with a signal logger and placed between Hamamatsu campus building for 24 hours. Then we plotted the data in Skyplot mode (Fig. 6) by RTKLIB. The data show GNSS constellations on the recorded location, which also find the visible satellites as in Fig. 7. In

the future, we planned to use the constellation data provided from governments that operate the GNSS in the proposed system.

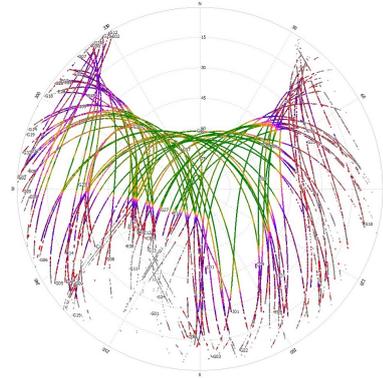


Figure 6: The observation data in Skyplot mode

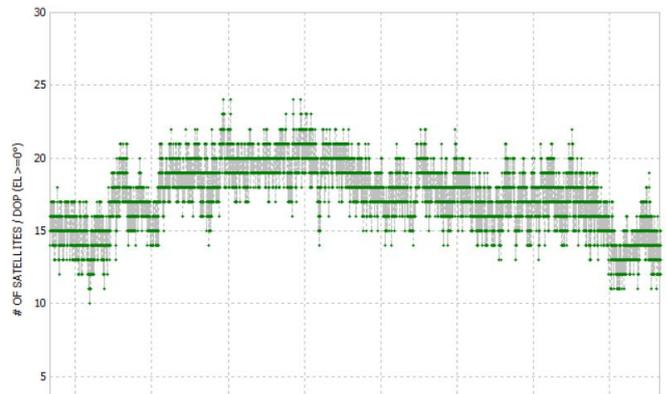


Figure 7: The number of the visible satellites from observation data

4.2.2 NLOS Detection by Obstacle Area Boundary

As mentioned in Sec. 3, the multipath signals from NLOS satellites is the potential problem that affect RTK accuracy. We simulated the NLOS satellites detection by using a fish-eye camera’s photo above the recorded location. The detection result should be same as using 3D map because of fisheye camera can observe surrounding obstacles above antenna[15]. As shown in Fig. 6, the difference color in the constellations represents an SNR of received signals. Best SNR values are plotted in green and the lower values are plotted in yellow, pink, blue, red and gray in order. We combined those data with a fisheye camera’s photo that took above the logging location (Fig. 8). As a result, SNR values started to drop significantly when navigation satellites arrived a tallest edge of a building.

Then, we did an elevation mask pattern by RTKLIB, for making an obstacle area boundary[6]. The result of elevation mask pattern is shown in Fig. 9.

From the result in Fig. 9, we can detect which visible satellite is NLOS. The signals that located out of the elevation mask pattern were transmitted by NLOS satellites.

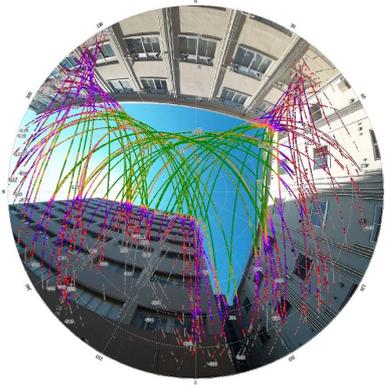


Figure 8: A fisheye photo combined with an observation data

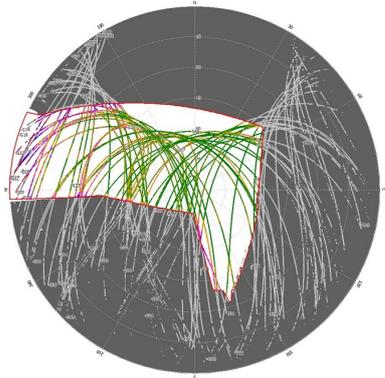


Figure 9: An observation data with elevation mask pattern

4.2.3 GDOP Calculation of LOS Satellites

From the previous procedure, we can detect and eliminate the NLOS satellites from observation data. This procedure will calculate a GDOP (Geometric dilution of precision) of the LOS satellites. The accuracy of RTK-GNSS is affected by several factors. One such factor is GDOP, which represents the geometric locations of the GNSS satellites as seen by the receiver. Better geometry will make a better position accuracy[16]. Meaning of DOP values are listed in Table 1.

We conducted a GDOP calculation from previous procedure data using RTKLIB. Figure 10 shows the GDOP values for 24 hours. Values were varied depend on LOS satellites geometry. We selected the example data from good and poor GDOP periods and plotted sky location of the LOS satellites as Figs. 11 and 12, orange circle is the LOS satellite and green circle is a receiver. Figure 11 shows a good alignment of the LOS satellites geometry, GDOP value is 2.3. In case of Fig. 12, the LOS satellites are aligned in a poor geometry,

Table 1: Meaning of DOP values

DOP	Ratings
1	Ideal
2 – 4	Excellent
4 – 6	Good
6 – 8	Moderate
8 – 20	Fair
20 – 50	Poor

GDOP value raised up to 40.

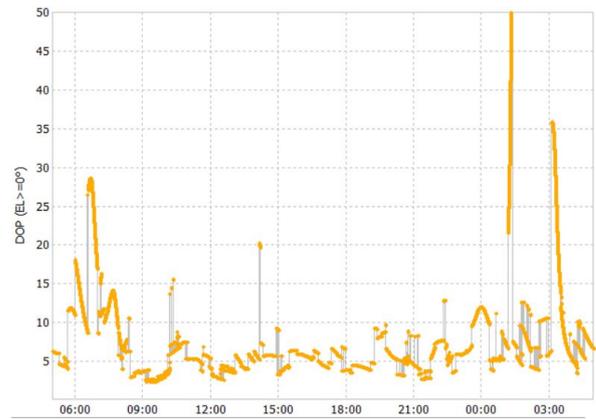


Figure 10: A 24 hours of GDOP values from the LOS satellites

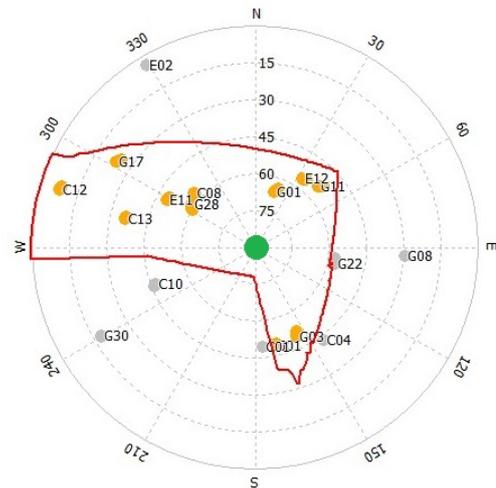


Figure 11: An example of LOS satellites geometry in good GDOP period

4.3 A Scoring Scheme for Forecasting Result

We introduced a scoring scheme for the forecast system evaluation result. We used three variables for calculation: quantity and GDOP of LOS satellites. We divided the poor, normal and good performance of each criteria, LOS satellites quantity and GDOP by data analysis from the recorded observation data, and additional GDOP table from the reference paper [16]. The first analyzed criteria are a quantity of LOS satellites. The data were recorded in an open sky area for two times, two hours each time. Then we used RTKLIB to analyze the data and compare the position results. FIX Solution, FLOAT solution, and no solution positioning are plotted in graph as Fig. 13. As mentioned earlier, minimum of four LOS satellites are required to calculate a receiver position, we can be observed this event when the quantity equal to 3, there are no positioning result. A higher number of the LOS satellites, let the positioning results better and eliminate “no

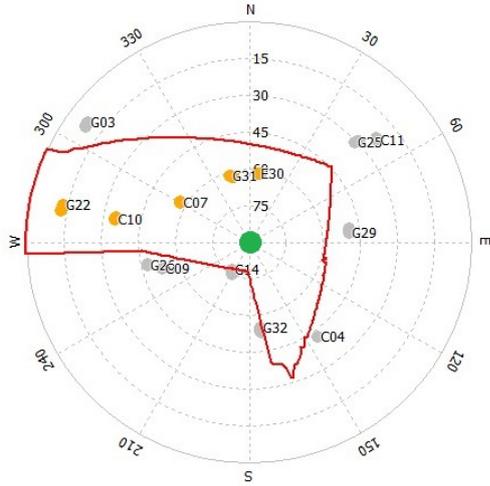


Figure 12: An example of LOS satellites geometry in poor GDOP period

solution” data. We set quantity of 4–6 in poor category, 7–11 in normal category and over 11 in good category.

The second analyzed criteria are a GDOP of LOS satellites. We have recorded an observation data with a signal logger and placed between Hamamatsu campus building on May 18, 2017, then we did a post processing on the recorded data. We founded that the GDOP 1–3 values inside the recorded data are not enough to measured. We assumed that these may cause from the surrounded obstacles. As an open sky data can be obtained a good GDOP, then we used the open sky data recorded from the previous criteria to measure the GDOP 1–3 values. The measurement data were combined and % FIX Solutions are plotted in graph as Fig. 14. As the result, trend of the % FIX solution continues to decrease when GDOP value getting higher. These results are also aligned with the reference paper [16]. According to the result, we set the GDOP values under 4 in a good category, 4–7 in a normal category and over 7 in a poor category.

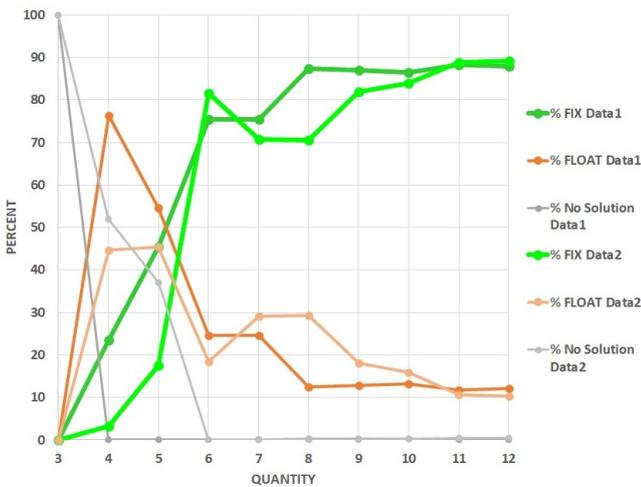


Figure 13: The graph of positioning results by a quantity of LOS satellites

Criteria of each variable are shown in Table 2. All crite-

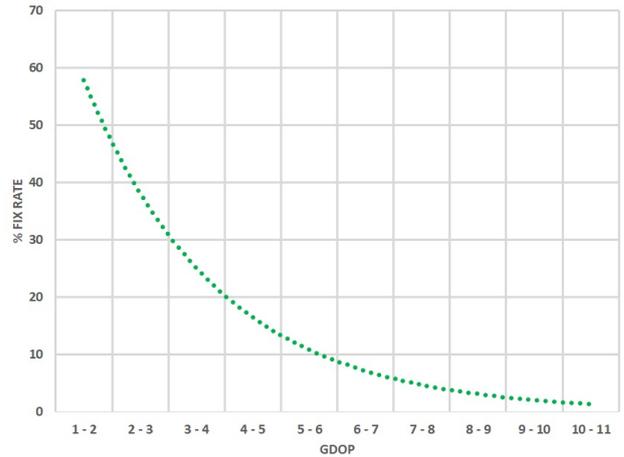


Figure 14: The graph of FIX positioning results by GDOP

Table 2: Criteria of each variable in the forecast system

		RTK Impossible	RTK Possible		
			Poor	Normal	Good
LOS Satellites	Quantity	< 4	4 - 6	7 - 11	> 11
	GDOP		> 7	4 - 7	< 4

ria have a relation to each other such as quantity of the LOS satellites will affect to the GDOP directly by its geometry. The best accuracy of an RTK navigation can be gained when all of variables are in the good criteria’s.

We converted the criteria above into a numbering score. Poor, normal and good criteria are “0”, “1” and “2” in order, for the quantity of LOS satellites and GDOP. Those variables are the primary affected to the RTK solution. Conversion table is shown in Table 3.

The final result is a summarize of two variables. The value will be varied from negative numbers to positive numbers. In case that the quantity of LOS satellites is under four from the beginning of the calculation, the system will output the result in “Not available” automatically because RTK navigation is totally impossible, by a minimum of four LOS satellites are required to estimate the receiver position and time.

The normal standard value is “2”, by summary of quantity and GDOP score in normal category. The value under “2” may represents an RTK navigation performance is under normal to poor and may not be available on that time and place. In contrast, the values more than “2” will represents an RTK navigation is available and can calculate a precious positioning on that time and place. Greater values mean better navigation performance.

Table 3: Criteria in scoring scheme

		RTK Impossible	RTK Possible		
			Poor	Normal	Good
LOS Satellites	Quantity	N/A	0	1	2
	GDOP		0	1	2

4.4 Evaluation

We used the recorded data from Sec. 4.2.1 as our scoring evaluation, then selected the test case periods from the recorded data. We calculated score of each period by our scoring scheme. The result of calculation is shown in Table 4.

From the results in Table 4, Case 1 which has the highest forecasting score can gained the highest %FIX rate of the RTK solutions, followed by Cases 2 and 3. In Case 4 which has the lowest score, RTK can be obtained only 1% of FIX solution. According to the results, RTK navigation may not work in some situation such as a receiver received a good quantity of LOS satellites but GDOP is poor, or all of it are in poor SNR. We planned to test this scoring scheme in the future experiment to validate its accuracy and practical in various situation.

4.5 Our Contributions

The proposed system is designed to be the pioneer system to forecast a precision by given time and position. This system introduces the idea that 3D Map and satellites constellation are not only improved real-time navigation or post-processing, but also can forecast which area is capable for high precision navigation.

5 CONCLUSION

In this paper, we have proposed a precision forecasting system. This system will expand a possibility of RTK-GNSS usage, by helps calculates and delivers which place is capable for high precision positioning. Through an experiment under the urban environment situation confirm that the results significantly improve in FIX solution and accuracy, by adjusting the parameters in RTKLIB configuration, respect to the NLOS detection result.

For the future work, a constellations data from a government will be used in the system, for more accurate in visible satellites detection. 3D map can be enhanced by the local observation data that obtained by the logger unit. The scoring scheme can be improved for more accurate result. We planned to test the proposed system in real situation with the forecasting score and validate overall performance.

ACKNOWLEDGEMENT

This work was supported by JSPS KAKENHI Grant Numbers 17H01731 (Grant-in-Aid for Scientific Research (B)).

REFERENCES

- [1] Czech Space Portal: GNSS - Global Navigation Satellite System, <http://www.czechspaceportal.cz/en/section-3/gnss-systems/>.
- [2] NovAtel: Real-time Kinematic (RTK), <http://www.novatel.com/an-introduction-to-gnss/chapter-5-resolving-errors/real-time-kinematic-rtk/>.
- [3] B. Hofmann-Wellenhof, H. Lichtenegger, and E. Wasle: "GNSS - Global Navigation Satellite Systems" First edition, *Springer-Verlag Wien*, 2008.
- [4] SCIENCING: What Is Difference Between RTK Fix & RTK Float? <http://sciencing.com/difference-between-rtk-fix-rtk-float-12245568.html>.
- [5] W.H. Kevin, G. Peter, and C.L. Don: "GPS accuracy part 2: RTK float versus RTK fixed," CREWES Research Report - Volume 22, 2010.
- [6] H. Tokura and N. Kubo: "Effective Satellite Selection Methods for RTKGNSS NLOS Exclusion in Dense Urban Environments," Proc. of the 29th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS+ 2016), pp. 304–312, 2016.
- [7] T. Kitani and H. Hatano: "A cooperative GPS/GNSS positioning method with neighboring receivers," Proc. of the International Workshop on Informatics (IWIN2014), pp. 145–153, 2014.
- [8] PennState: Relative Positioning, <https://www.e-education.psu.edu/geog862/node/1725>.
- [9] Trimble: Critical factors affecting RTK accuracy, http://www.trimble.com/OEM_ReceiverHelp/V4.44/en/PositionModes_CriticalFactorsRTK.html.
- [10] I.G. Petrovski: "GPS, GLONASS, Galileo, and Beidou for Mobile Devices" First edition, *Cambridge University Press*, 2014.
- [11] W. Lei, D.G. Paul, and K.Z. Marek: "GNSS Shadow Matching: Improving Urban Positioning Accuracy Using a 3D City Model with Optimized Visibility Prediction Scoring," Proc. of the 25th International Technical Meeting of the Satellite Division of The Institute of Navigation, Nashville TN, pp. 423–437, 2012.
- [12] It's Time for 3D Mapping-Aided GNSS, September/October 2016 <http://www.insidegnss.com/node/5096>, 2016.
- [13] A. Kumar, Y. Sato, T. Oishi, and K. Ikeuchi, "Identifying Reflected GPS Signals and Improving Position Estimation Using 3D Map Simultaneously Built with Laser Range Scanner," 13th ITS Asia-Pacific Forum, Auckland, New Zealand, April 2014.
- [14] T. Takasu and A. Yasuda: "Development of the low-cost RTK-GPS receiver with an open source program package RTKLIB," Proc. international symposium on GPS/GNSS, pp. 4–6, 2009.
- [15] M. Julien, A. Sebastien, and R. Yassine: "Fisheye-Based Method for GPS Localization Improvement in Unknown Semi-Obstructed Areas," *Sensors* 2017, 17(1), 119, <http://doi.org/10.3390/s17010119>, 2017.
- [16] V.B.S.Srilatha Indira Dutt, G.Sasi Bhushana Rao, S.Swapna Rani, Swarna Ravindra Babu, Rajkumar Goswami and Ch.Usha Kumari: "Investigation of GDOP for Precise user Position Computation with all Satellites in view and Optimum four Satellite Configurations," *J. Ind. Geophys. Union* (July 2009), Vol. 13,

Table 4: Test case data with scoring results

	Date & Time	LOS Satellites		Predicted quality	% Fix rate	% Float rate	% No solution
		Quantity	GDOP				
Case 1	July 14, 2017 01:30 - 03:30 GPST	31	1	4	70.55	28.33	1.10
Case 2	May 18, 2017 9:30 - 9:35 GPST	11	2.5	3	2.81	97.19	0
Case 3	May 18, 2017 10:40 - 10:50 GPST	8	7	2	6.94	93.06	0
Case 4	May 19, 2017 01:15 - 01:20 GPST	5	31	0	1.00	16.61	82.39

No. 3, pp. 139–148, 2009.

- [17] A.T. Irish, D. Iland, J.T. Isaacs, J.P. Hespanha, E.M. Belding, and U. Madhow: “Using crowdsourced satellite SNR measurements for 3D mapping and real-time GNSS positioning improvement,” S3 ’14 Proceedings of the 6th annual workshop on Wireless of the students, by the students, for the students, pp. 5–8, 2014.
- [18] H. Li-Ta, C. Feiyu, and S. Kamijo: “Evaluation of Multi-GNSSs and GPS with 3D Map Methods for Pedestrian Positioning in an Urban Canyon Environment,” IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences Vol. E98-A, No. 1, pp. 284–293, 2015.
- [19] M. Obst, S. Bauer, P. Reisdorf, and G. Wanielik: “Multi-path Detection with 3D Digital Maps for Robust Multi-Constellation GNSS/INS Vehicle Localization in Urban Areas,” Proc. 2012 IEEE Intelligent Vehicles Symposium (IV), pp. 184–190, 2012.
- [20] GISLab, Los Alamos National Laboratory: GPS Guide, http://gislab.lanl.gov/gps_guide.html.

Session 6:
Systems and Applications
(Chair: Kozo Okano)

Statistical Prediction of Wagyu Characteristics with Bayesian Hierarchical Modeling

Hisami Sano^{†*}, Takatoshi Fujiki[†], Takuya Yoshihiro^{†**},
Haruka Ikegami[§], Tamako Matsuhashi[§], and Kazuya Matsumoto[§]

[†]Graduate School of Systems Engineering, Wakayama University, Japan

[‡]Faculty of Systems Engineering, Wakayama University, Japan

[§] Graduate School of Biology-Oriented Science and Technology, Kinki University, Japan
{*s161025, **tac}@sys.wakayama-u.ac.jp

Abstract - Japanese brand Wagyu is popular as expensive and excellent foodstuff not only in Japan but also all over the world. Developing a methodology to efficiently produce high-value beef cattle is an economically important research goal that satisfies the large demands in both Japan and the world. In genetic inbreeding, we have a method to produce genetically excellent sire that produces high-value beef cattle, and it is practically carried out. However, in raising and fattening cattle, we do not have any scientifically proved methodology to improve economical traits of quality beef cattle. In this paper, we propose a method to predict economical traits of quality beef from various values measured periodically during the fattening period. As the measured values, we not only suppose basic growth indices such as body height and weight but also include analytical measurements such as blood components and gene expressions. We apply Bayesian hierarchical modeling to predict economic traits. Within the framework of Bayesian hierarchical modeling, we propose a statistical model that is suitable to predict economical traits of brand beef. Through evaluation with real data, we show that our model has better prediction ability than the multiple regression analysis, which is a general multivariate prediction model used in this area.

Keywords: Wagyu, Carcass Characteristics, Bayesian hierarchical modeling

1 INTRODUCTION

Japanese brand Wagyu is popular as expensive and excellent foodstuff not only in Japan but also all over the world. Developing a methodology to efficiently produce high-quality brand Wagyu is an economically important research goal that satisfies the large demands in both Japan and the world.

Wagyu is a high-quality brand beef that is generated in several specific areas in Japan, which is produced by crossbreeding the genetically-excellent sires [1], [2]. Generally, calves are generated from frozen sperm of sires, and are sold as beef after the breeding period of 2-3 years with fattening farmers. Beef is rated and evaluated with 6 indices, and the price is determined. The rating indices, which we call the economical traits hereafter, are carcass weight, BMS (Beef Marbling Standard), yield enhancement, rib thickness, subcutaneous fat thickness, and rib-eye area [3]. Especially, carcass weight, which refers the total beef weight, and BMS, the marbling

level, are regarded as the most important indices that determine the economical value of beef.

There have been various efforts to produce high-value beef cattle effectively. First, breeds of sires have been improved from a genetic perspective. Specifically, by selecting genetically excellent individuals to produce the next-generation sires repeatedly, the genetic ability of sires to produce high-value beef has been improved so far. To evaluate the genetic ability of sires, an index called "breeding value" is generally referred. A breeding value of a sire is a reference value that represents the genetic ability on a particular economical trait, which is statistically calculated from the carcass characteristics scores (i.e., economical traits) of its descendant calves in consideration of their blood relationship. As the method to estimate the breeding value, BLUP method [4] is known and widely used on the sites of brand beef. Although the genetic mechanism in expressing economical traits has not been clarified at all, the breeding value is known to inherit significantly in statics. Consequently, by selecting better individuals that has improved breeding values in breeding sires, we can improve the ability of sires in terms of carcass characteristics.

On the other hand, considerable efforts in raising beef cattle have been dedicated to improve the economical traits of quality beef. Fattening farmers buy calves from breeding farmers and bring them up to produce high-value beef cattle with their own know-how. In general, fattening farmers control feeds, environment, stresses of cattle during fattening to raise high-value beef cattle. For example of such kind of controls, it is empirically known that the beef quality can be improved by controlling vitamin-A concentration, in other words, controlling feeding of cattle during the fattening period [5]. However, because the way of fattening (e.g., feeding) is based on the farmers' know-how and experience, it is basically difficult to share or inherit the way of fattening cattle. It is important to establish a methodology of fattening based on data, i.e., from the viewpoint of science and technology, in order to produce quality beef efficiently. It is strongly desired to develop a new fattening methodology to produce high-value beef cattle based on data science.

In this paper, we propose a method to predict economical traits of quality beef from various values measured periodically during the fattening period. As the measured values, we not only suppose basic growth indices such as body height and weight but also include analytical measurements such as blood components and gene expressions. Note that if the eco-

conomic traits can be predicted from the measured values, we would conversely utilize the methodology to improve the economic traits by controlling the measured values to meet their optimal values. In this study, we apply Bayesian hierarchical modeling [6] to predict economic traits. Bayesian hierarchical model is a mathematical model that models the target values by multiple-levels statistic distributions with relatively a small number of parameter values. With this method, we can incorporate multiple factors that influence on the target value into the model, and obtain the optimal probabilistic distribution that explains the data. From those characteristics, the hierarchical Bayesian model is suitable for the cases in which the number of samples is small and we would like to treat individual differences or additive random effects including our case in this study, i.e., in which we would like to build a model to explain economical traits using raising data. However, to the best of our knowledge, there is no existing work that apply hierarchical Bayesian model and treat additive random effects in the study of livestock. In this paper, we propose a new specific model to explain economical traits of Wagyu within the framework of hierarchical Bayesian modeling, in which we incorporate random effects to the basic multiple regression analysis.

The reminder of this paper is organized as follows. In Sec.2 we explain the basic knowledge of Wagyu. In Sec.3, We explain about Bayesian modeling and its Stratification. In Sec. 4, we describe the proposed method, and present the evaluation results in Sec.5. Finally, we conclude the work in Sec.6.

2 Wagyu: A Brand Beef of Japan

Japanese Black Cattle is a beef cattle peculiar to Japan, which produces various brand beef called Wagyu such as Kobe beef, etc. There are many regional brand beefs in Japan, each of which has its own way to breed cattle, and apply its own criterion to authorize whether each head of cattle is sold under the name of the brand beef. As the authorization criteria, there are several items, e.g., the birth of cattle, the way to raise cattle, the rating of beef, etc. Among them, the rating is the most important. The rating criteria include various values, and especially 6 items among them are regarded as the most important ones to judge whether a head of cattle is authorized as brand beef. The 6 items, which we call economical traits, are carcass weight, BMS (beef marbling standard), yield enhancement, rib thickness, subcutaneous fat thickness, and rib-eye area. From these criteria, the price of beef in the market is determined. Therefore, the farmers of brand beef have been made a great endeavor to produce quality beef.

Wagyu farmers take various methodologies to produce quality beef stably. One of the most important methods is to control bloodline so as to have better values of the economical traits. Since the bloodline is known to have close relationship with economical traits, efficient inbreeding by producing and identifying genetically excellent individual cattle has a significant importance to improve the value of brand beef site. Each brand-beef site usually breeds several head of cattle called sires that have excellent genetic ability. From sires, we take sperms and freeze them, and sell them to farmers. With this system, excellent bloodline of sires is distributed to farm-

ers and generate thousands of children cattle from a excellent sire. Note that, in brand-beef sites, father of each beef cattle is called '1-generation ancestor' and the father of beef cattle is one of the most important criteria to predict economical traits of beef cattle.

As a statistical methodology to predict economical traits of beef cattle from past records, the breeding values are usually used in brand beef sites. The breeding values are calculated for each economical trait, which represent the ability to improve the trait values compared to the average ability in the group. There are two kinds of breeding values, i.e., estimated breeding values and expected breeding values. The former is calculated for sires who have descendants with carcass characteristic scores and represents the ability to improve 6 economical traits. In contrast, the latter is calculated for each beef cattle that does not have enough number of descendants to estimate breeding values.

We have several variations of bloodline models used to compute breeding values. Currently, the most frequently used model is so called 'animal model,' which considers all the relative relationship including brothers of beef cattle that have the same mother. With a bloodline model and the data set, BLUP method calculates the breeding values in a statistical manner as the genetic ability inherited through bloodlines [4]. The expected breeding value for each beef cattle is calculated as the average of its two parents.

On the other side, raising method to produce high-value beef cattle stably also has been studied so far. However, methods in this area are mostly depends on experiences of farmers, and are not based on any scientific results or real data. For example, livestock associations or stock farmers have accumulated their experience to raise high-value beef cattle as know-how or some kind of manuals. This kind of information has wide variations from direct methods such as how to feed cattle to indirect methods such as the structure of cowsheds. As for the academic results, a few studies have been published on the relationship between raising methodology and economical traits. For example, there is a study on improving BMS values by controlling the concentration of vitamin A [5]. However, in the current state, we have still too little knowledge to actually control economical traits in raising in livestock farms. In this paper, we try to model economical traits with Bayesian hierarchical model to predict them from a dataset of time-series measurements in raising cattle.

3 HIERARCHICAL BAYESIAN MODEL

The hierarchical Bayesian model is a statistical model in which the probability distribution of parameters has a hierarchical structure. In the hierarchical Bayesian model, it is possible to construct a statistical model that fits the given data set by computing an optimal prior probability distribution using Bayes' theorem.

In this section, we will explain the Bayes' theorem, which is the basis of the hierarchical Bayesian model, in Section 3.1. Based on this, we will explain the hierarchical Bayesian model in Section 3.2 section.

3.1 Bayes' Theorem

Bayes' theorem is a theorem indicates a property on conditional probability. In Bayesian statistics, Bayes' theorem is expressed by the following equation,

$$P(H|D) \propto P(D|H)P(H),$$

where H denotes a hypothesis, D does a given observation data, $P(H)$ does a prior probability distribution, $P(H|D)$ does the posterior probability distribution, and $P(D|H)$ does the likelihood, respectively. The prior probability distribution is defined as the probability that the hypothesis H holds, and the posterior probability distribution is the probability that the hypothesis H holds when data D is given. The likelihood is defined as the probability to obtain data D when the hypothesis H holds.

From this equation, it is clearly shown that a posterior probability distribution with higher objectivity can be obtained by multiplying the prior probability distribution, which is the subjective probability distribution, by the likelihood that includes objective facts. This kind of methodology to estimate the posterior probability distribution from prior probability distribution and the likelihood is called Bayesian inference. In addition, it is possible to calculate a posterior probability distribution with higher objectivity by regarding this posterior as a prior and apply the likelihood obtained by a new data set. The procedure to update the posterior probability distribution as above is called Bayes updating.

3.2 Bayesian Models

The hierarchical Bayesian model is a class of statistical models in which the parameters of the probability distributions have a hierarchical structure in the Bayesian statistical model. Namely, by expressing the parameters of probability distributions by other probability distributions, a hierarchical statistical model is constructed. If we consider the probability distributions of parameters in a hierarchical statistical model and consider them as prior probability distributions, we call them a hierarchical prior probability distribution and we call their parameters hyper parameters. Also, when considering the probability distributions of the hyper parameters as prior distributions, we call them hyper prior probability distributions. The structure of hierarchical distributions is effective to introduce the parameters that represent random effects in a model. The random effect is the effect that explains the variance among individuals or groups while it does not change the variance of the whole data set, which is known to occur frequently in the real world. The hierarchical Bayesian model has a flexibility to afford variations of parameter distributions by allowing wide-range of hyper parameter values in fitting to a given data set, while unchanging the hierarchical prior probability distribution common to the whole data set. This property allows us to explain the nature of random effects that tend to make a set of data values excessively scattered. From these properties, the hierarchical Bayesian model is capable of constructing a model that accurately explains actual data with few parameters.

In general, parameters of the hierarchical Bayesian model are optimized by MCMC (Markov Chain Monte Carlo) method, which is an optimization algorithm based on randomized values. The MCMC method makes parameters converge to the values that represent the optimal probability distribution by repeating probabilistic sampling based on random numbers, and it obtains the posterior probability distribution with the optimal parameters and the hyper parameters under the input data.

4 PROPOSED METHOD

4.1 Overview

In this paper, we build a model that explains the carcass characteristics of Wagyu under the framework of hierarchical Bayese modeling. In Wagyu, since the labor to raise cattle is significantly large, it is generally difficult to collect a large number of samples in the data set. For this kind of small data set, multiple regression analysis is often applied that expresses the objective variable as the linear sum of several explanatory variables. The analysis of carcass characteristics of Wagyu also meets the suitable cases to apply regression analysis. However, since the carcass characteristics are under effects of the environment, e.g., the way of raising by farmers, it is desirable to apply more flexible models that enables us to incorporate environmental items in the model. As one of the flexible models that suffice the requirements, we utilize hierarchical Bayesian model, which hierarchically defines the parameters of linear models. The hierarchical Bayesian model explains the objective variable with relatively a small number of explanatory variable by defining hierarchically the statistical distributions. In the framework of hierarchical Bayesian modeling, we design a statistical model in which we express the fixed effects with time-series raising items, and express random effects with environmental items. By incorporating environmental effects to the model, our model achieves more accurate prediction of carcass characteristics of Wagyu. So, in this paper, we construct a hierarchical Bayesian model, which is possible to treat both fixed effects and random effects, for predicting carcass characteristics of beef cattle. In the following, we first describe the data structure and the format assumed in our method in Section 4.2. We next describe the specific model of the proposed method in Section 4.3.

4.2 Data Format

The input data set is classified into 3 parts. We call the first one raising data, the second environmental data, and the last one economical traits data.

The raising data represents various values measured during the fattening period. Various data items can be included as the measured items. For example, in addition to basic items such as body height and weight of cattle, we can also consider the amount fed, the value of blood constituents, etc. These raising items could be measured over time during the raising period of cattle. We let the set of measurement items measured during the raising period be X . In X , we not only includes the basic items such as weight and body height, but also con-

sider some specific items such as the amount of vitamin A in serum or expression values of genes or proteins etc. Note that, although several items may be measured in time series at a certain interval, we in this paper do not consider the aspect of time since the time interval is not sufficiently large in this data set. We let B be the set of cattle in the raising data, and let $b \in B$ be each individual cattle. Also, we let $v_{b,x}$ be the value of raising item $x \in X$ for cattle b .

The environmental data include the parameters in the environment in which each cattle was raised up. In this paper, it is mainly used to express variables to consider group effects. As candidates for the environmental data items, we can consider region where cattle are grown up, season of birth, sires, type of food, farmers, etc. In this paper, we let E be the set of raising environments, and if the cattle $b \in B$ belongs to environment $e \in E$, then $w_{b,e} = 1$, and $w_{b,B,e} = 0$ otherwise.

The economical traits data is a set of numerical values that represents the carcass characteristics scores for each individual cattle. There are many data items that represents carcass characteristics, e.g., SFT represents the thickness of subcutaneous fat, RT does rib thickness, and BMS does grade of marbling, etc. We denote such a set of items to rate carcass characteristics by P , and we let p_b be the value of item $p \in P$ for individual cattle b .

4.3 The Proposed Models

In this paper, we propose a Bayesian hierarchical model to include random effects of environmental items. In these models, we regard raising items and environmental items as explanatory variables, and carcass characteristics items as objective variables. In this paper, it is assumed that the carcass characteristic score p_b of cattle b follows the normal distribution $N[\mu_b, \sigma]$ with mean μ_b and standard deviation σ . Then, we use a noninformative prior distributions as the prior probability distribution of σ because the standard deviation σ is hardly determined beforehand. A noninformative prior is a prior probability distribution such that it is designed not to affect the posterior probability distribution as much as possible.

In the proposed model, the average μ_b is expressed as an extension of multiple regression equation. Namely, μ_b is expressed by a linear sum of the multiple regression terms and the random effects of environmental items. We show the formula of the model in the following, with the model structure in Fig. 1.

$$\mu_b = M + \sum_{i \in I} \beta_i v_{b,i} + \sum_{e \in E} \gamma_e w_{b,e}. \quad (1)$$

In this equation, we let M be the value representing the intercept in the multiple regression equation. We let I be the set of items selected from X , i.e., $I \subseteq X$, as explanatory variables. We let β_i be the regression coefficient of raising item i . We let γ_e be the random effect of environment $e \in E$, which shows how much difference exists compared to the average if cattle b belongs to e . Also, we let $w_{b,e}$ be the variable that takes 1 if cattle b belongs to e and takes 0 otherwise. Note that, in this model, you can treat multiple groups of environmental items, e.g., both random effects of farmers and sires. In this case, you can simply replace E with E_1, E_2, \dots ,

where E_1 is a set of environmental items on farmers and E_2 on sires.

Since the coefficients β_i and intercept M cannot be determined arbitrarily, we use a noninformative prior as a prior probability distribution. Since the random effect γ_e represents how much the average of the carcass characteristics scores is different compared to the average if cattle b belongs to environment e , we assume that the prior probability distribution of γ_e follows the normal distribution $N[0, \sigma']$ where 0 is the average and σ' is the standard deviation. Here, since the standard deviation σ' cannot be determined arbitrarily, we use the noninformative prior as the hyper prior probability distribution of σ' .

5 EVALUATION

5.1 Method

We evaluate the performance of the proposed model using real data. This set of data includes measured values obtained from 90 head of brand cattle fattened in 3 regions.

This data set includes expression levels of multiple proteins in the serum taken from each individual cattle three times during the whole fattening period, say, in the initial period, the middle period, and the late period. We selected 6 proteins, which we expect to correlate with carcass characteristics scores from the results of preliminary experiments. Namely, the data set contains 18 protein expression levels per individual cattle.

As environmental items, we used three items: grown-up regions, sires and farmers, where sires here mean the farther of each individual cattle. Each environmental item has 3, 5, and 8 groups, and each individual cattle is included in one group corresponding to each environmental item.

As the data items on carcass characteristics, the data includes all of 6 major economical traits in the evaluation of brand beef. Specifically, they are carcass weight(CW), beef marbling standard(BMS), yield enhancement(YE), rib thickness(RT), subcutaneous fat thickness(SFT), and rib eye area(REA), where YE is the reference value that express the amount of beef yielded from an individual cattle.

We implemented the proposed method using R and Rstan, and applied it to the data. In the proposed model, we limited the number of items used for multiple regression is limited to 2, i.e., $I = 2$, in order to reduce the number of free parameters in the model. We executed the parameter optimization for all combinations of 2 items from 18 items and used the best one in our evaluation result. As environmental items (i.e., groups), we tried every combination of 1 or 2 items from the three items in the data set, i.e., fattening regions, sires and farmers. We made evaluations independently for each of the 6 items on carcass characteristics scores. We compute the optimal parameter values using MCMC method for all of these cases mentioned above, and regard the expected value μ_b of each carcass scores item. To evaluate the prediction accuracy with the proposed model, we performed 5-division cross-validation for all cases, i.e., all combinations of carcass scores items and environmental items. We used the Mean Absolute Error (MAE) as the evaluation criterion.

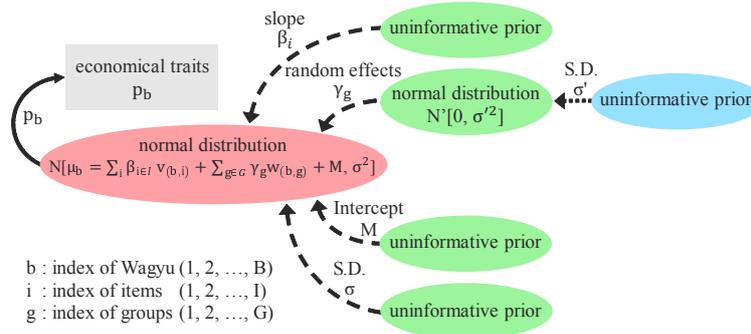


Figure 1: The First Proposed Model (Including Random Effects in Intercept)

MAE is the mean value of the absolute errors E_{p_b} represented by

$$E_{p_b} = |Real(p_b) - Pred(p_b)|$$

where $Real(p_b)$ is the real value of p_b in the data set, and $Pred(p_b)$ is the predicted value of p_b computed in the proposed method.

As the conventional method to compare with, we used the multiple regression analysis, which does not include the random effect. In the multiple regression analysis, prediction value for each cattle was calculated based on the regression line, i.e., the value on the line is the prediction value, and MAE was calculated by performing 5-division cross validation, which is same as the proposed method. Also, in the multiple regression analysis, we tried the cases of 2, 5, and 9 items, for each of which we computed all combinations of all 18 items and used the best MAE value for the comparison. When comparing these with the proposed method, we use the best MAE value among the three cases (i.e., 2, 5, and 9 items) as the representative MAE value of multiple regression analysis.

5.2 Evaluation Results

5.2.1 Results of Multiple Regression Analysis

Figure 2 shows the results on multiple regression analysis for each number of items. This result shows the best MAE value calculated in multiple regression analysis with three cases, i.e., with 2, 5, and 9 items. As the number of items, three results of 2 items, 5 items, 9 items were shown. From this result, we can see that there is no significant improvement of MAE values even if we increase the number of items. This shows that the linear regression itself has a limitation on explaining economical traits on carcass characteristics, and that the linear regression is not sufficiently effective as a model to predict carcass characteristics scores, even if we apply variables to explain it.

Regarding individual traits on carcass characteristics, we see that carcass weight (CW) and rib eye area (REA) are improved by increasing the number of items; The MAE of carcass weight improved about 2 [kg], and that of rib eye area improved about 0.5[cm], when the number of items increased from 2 to 5.

MAE of Multiple regression analysis						
	CW (kg)	YE	REA (cm ²)	RT (cm)	SFT (cm)	BMS
2 items	34.031	2.6	6.928	0.563	0.537	1.583
5 items	31.651	2.535	6.493	0.641	0.576	1.577
9 items	31.728	3.162	6.473	0.641	0.622	1.672

Figure 2: Results with Multiple Regression Analysis

5.2.2 Proposed Method

The evaluation result of the proposed model (1) is shown in Fig. 3. The number of groups in the figure means the number of environmental items, and we compare the results for each environmental item and each pair of the items considered as random effects. As mentioned earlier, ‘MAE (proposed)’ shows the best result from all combinations of every two items in the regression part, and ‘MAE (MRA)’ where MRA means Multiple Regression Analysis, shows the best result of MAE in multiple Regression Analysis of 2, 5 and 9 items shown in the previous section. ‘Diff.’ is the value obtained by subtracting the MAE of multiple regression analysis from that of the proposed method, meaning that the lower the value, the better the proposed method is. Avg. and S.D. represent the mean and standard deviation of the trait values. Also, in order to estimate the degree of prediction accuracy, we show the relative amount of errors compared with the standard deviation, i.e., Diff. divided by S.D.

From this result, the average of MAE with the proposed model (1) is better than that of multiple regression analysis in almost all traits except the rib-eye area, which shows that the proposed model improves the prediction accuracy by applying random effects of environmental items. In the five traits that improved the prediction accuracy, the improvement compared with the standard deviation, i.e., Diff. / S.D., is about a few percent. Focusing on the number of free parameters, MAE of the proposed method, which has 8 free parameters when we applied one environmental item ‘sire,’ is smaller than the result of multiple regression analysis with 9 items. Since the prediction accuracy is low while the number of parameters is almost the same, we can understand that the improvement of accuracy in the proposed method is not caused

CW (kg)							
# of groups	groups	MAE (proposed)	MAE (MRA)	Diff.	Avg.	S.D.	Diff./S.D.
1	region	31.82	31.651	0.169	476.41	43.171	0.004
	sire	30.14		-1.511			-0.035
	farmer	30.46		-1.191			-0.028
2	region and sire	29.347		-2.304			-0.053
	region and farmer	29.75		-1.901			-0.044
	sire and farmer	29.434		-2.217			-0.051

(a) carcass weight

YE							
# of groups	groups	MAE (proposed)	MAE (MRA)	Diff.	Avg.	S.D.	Diff./S.D.
1	region	1.024	2.535	-1.511	69.287	5.446	-0.277
	sire	1.056		-1.479			-0.272
	farmer	1.045		-1.49			-0.274
2	region and sire	1.075		-1.46			-0.268
	region and farmer	1.049		-1.486			-0.273
	sire and farmer	1.054		-1.481			-0.272

(b) yield enhancement

REA (cm ²)							
# of groups	groups	MAE (proposed)	MAE (MRA)	Diff.	Avg.	S.D.	Diff./S.D.
1	region	6.67	6.473	0.197	59.722	8.902	0.022
	sire	6.511		0.038			0.004
	farmer	6.774		0.301			0.034
2	region and sire	6.813		0.34			0.038
	region and farmer	7.052		0.579			0.065
	sire and farmer	6.841		0.368			0.041

(c) rib eye area

RT (cm)							
# of groups	groups	MAE (proposed)	MAE (MRA)	Diff.	Avg.	S.D.	Diff./S.D.
1	region	0.556	0.563	-0.007	8.546	0.712	-0.01
	sire	0.549		-0.014			-0.02
	farmer	0.52		-0.043			-0.06
2	region and sire	0.548		-0.015			-0.021
	region and farmer	0.513		-0.05			-0.07
	sire and farmer	0.517		-0.046			-0.065

(d) rib thickness

SFT (cm)							
# of groups	groups	MAE (proposed)	MAE (MRA)	Diff.	Avg.	S.D.	Diff./S.D.
1	region	0.532	0.537	-0.005	2.591	0.642	-0.008
	sire	0.546		0.009			0.014
	farmer	0.526		-0.011			-0.017
2	region and sire	0.551		0.014			0.022
	region and farmer	0.514		-0.023			-0.036
	sire and farmer	0.503		-0.034			-0.053

(e) subcutaneous fat thickness

BMS							
# of groups	groups	MAE (proposed)	MAE (MRA)	Diff.	Avg.	S.D.	Diff./S.D.
1	region	1.565	1.583	-0.018	7.4	2.102	-0.009
	sire	1.61		0.027			0.013
	farmer	1.601		0.018			0.009
	2	region and sire		1.57			-0.013
region and farmer		1.578		-0.005			-0.002
sire and farmer		1.595		0.012			0.006

(f) BMS

Figure 3: Results with The First Proposed Model

by the increase of free parameters, but by incorporating random effects under the Bayesian hierarchical model.

If we see each economical traits, (b) yield enhancement improved the most, which is about 27% with any group combination. In (a) carcass weight, the effect of sires is the largest and is followed by the effects of farmers and regions. Because the bloodline is known to be important for yielding quality beef of brand cattle, the result that the economical traits is largely influenced by sires (i.e., bloodline) is a reasonable tendency supported by the experiences of this area of study. On the other hand, in (f) rib thickness and (e) subcutaneous fat thickness, the effects of farmers are observed to be large. (f) BMS, which is as important economical trait as carcass weight in brand beef, has a large effect of regions. However, because farmers as well as bloodlines are different according to regions, this would be an indirect effect in which the root causes are still hidden within the effect of regions.

6 CONCLUSION

In this paper, we proposed a Bayesian hierarchical model that explains economical traits of brand beef using several measured values in the raising period of beef cattle. The proposed model incorporates environmental items to consider group effects as the slope and the intercept of multiple regression models. Our evaluation results show that the proposed model improves the prediction accuracy without increasing the number of free parameters compared to the conventional multiple regression analysis, meaning that considering group

effects is effective to explain economical traits of brand beef.

ACKNOWLEDGMENT

This work was partly supported by “the Program for Promotion of Stockbreeding” of JRA (Japan Racing Association).

REFERENCES

- [1] Wagyu Registry Association, “Compliation of Sires of Japanese Black Cattle,” 2003 (In Japanese).
- [2] A Guide of Japanese Black Cattle Sires, <http://liaj.lin.gr.jp/index.php/detail/data/m/803237096> (referred in May 2017) (In Japanese).
- [3] Japan Meat Grading Association, “The Manual of Standard in Dealing Pork and Beef,” 2001 (In Japanese).
- [4] N. D. Cameron, “Selection Indices and Prediction of Genetic Merit in Animal Breeding,” CAB International (1997).
- [5] Oka, A., Dohgo, T., Juen, M., and Saito, T., “Effects of vitamin A on beef quality, weight gain, and serum concentration of thyroid hormones, insulin-like growth factor-I, and insulin in Japanese black steers, Animal Science and Technology”(1998).
- [6] Takumi Kubo, “Introduction to Statistical Modeling for Data Analysis,” Iwanami Shoten Publishers, 2012 (In Japanese).

Best-time estimation method by region and tourist spot using information interpolation

Masaki Endo*, Masaharu Hirota**, Shigeyoshi Ohno*, and Hiroshi Ishikawa***

*Division of Core Manufacturing, Polytechnic University, Japan
{endou, ohno}@uitec.ac.jp

**Department of Information Science, Okayama University of Science, Japan
hirota@mis.ous.ac.jp

***Graduate School of System Design, Tokyo Metropolitan University, Japan
ishikawa-hiroshi@tmu.ac.jp

Abstract –Various studies have been conducted to analyze social media data in real time and to extract events in the real world. A benefit of analysis using data with position information is that it can accurately extract an event from a target area to be analyzed. However, because data with position information are scarce among all social media data, the amount to analyze is insufficient in almost all areas. In other words, the problem indicates that we cannot fully extract most events. Therefore, efficient analytical methods are necessary for accurate extraction of events with position information, even in areas with few data. For this study, we estimate the time of biological season observation in each area and sightseeing spot by information interpolation using tweet location information. Herein, we explain the analysis results obtained using interpolation information.

Keywords: trend estimation; phenological observation; Twitter; information interpolation

1 INTRODUCTION

In recent years, sightseeing has come to be regarded as an extremely important growth field in Japan to revive its powerful economy [1]. Tourism with a large economic ripple effect is expected to benefit from regional revitalization and employment opportunities by accommodating world tourism demand, including that from rapidly growing Asia. In addition, people around the world can discover and disseminate the charm of Japan and can promote mutual understanding with other countries.

In addition to the promotion of tourism to Japan, the progress of domestic travel is important. It is necessary for a nation with modern tourism to build a community society where regional economies are well-served, attracting tourists widely. Moreover, it is necessary to cultivate tourist areas full of individuality and to promote their charm positively.

According to a survey study of IT tourism and services to attract customers [2] by the Ministry of Economy, Trade and Industry (METI), tourists want real-time information and local unique seasonal information posted on websites. Current websites provide similar information in the form of guidebooks. Nevertheless, the information update frequency of that medium is low. Because each local government, tourism association, and travel company independently provides information about travel destination local unit, it is

difficult for tourists to collect information for “now” tourist spots. Therefore, providing current, useful, real-world information for travelers by capturing the change of information in accordance with the season and time zone of the tourism region is important for the travel industry.

We consider a method to estimate the best time for phenological observations for tourism such as the best time for viewing cherry blossoms and autumn leaves in each region by particularly addressing phenology observations assumed for “now” in the real world. We define “now” as information for tourism and disaster prevention required by travelers during travel, such as best flower-viewing times, festivals, and locally heavy rains.

Tourist information for best times requires a peak period, which means that the best time is not a period after or before falling flowers, but a period to view blooming flowers. Furthermore, the best times differ among regions and locations. Therefore, it is necessary to estimate the best time of phenological observation for each region and location. Estimating the best-time viewing requires collection of large amounts of information with real-time properties.

For this study, we use Twitter data obtained from many users throughout Japan. We use Twitter [3], a typical microblogging service, and also use geotagged tweets that include position information sent in Japan to ascertain the best time (peak period) in biological season observation by region. We proposed a low-cost estimation method [4]. Using this method, prefectures and municipalities showing a certain number of tweets with geotags can be estimated with a relevance rate of about 80% compared to the flowering day / full bloom day of cherry blossoms observed by the Japan Meteorological Agency. The geotagged tweets that are used with this method are useful as social indicators that reflect the real world situation. They are a useful resource supporting a real-time regional tourist information system in the tourism field. Therefore, our proposed method might be an effective means of estimating the best time to view events other than biological seasonal observations.

However, geotagged tweets are extremely few among all tweets. Therefore, a problem exists by which the amount of data is insufficient for analysis with finer granularity. For this reason, it is necessary to improve the method of interpolating the information of geotagged tweets to conduct further detailed analyses in areas such as sightseeing spots. For this research, we propose a method of estimating the

best time by tourist spot by performing information interpolation based on amounts of regional information.

The remainder of the paper is organized as follows. Chapter 2 presents earlier research related to this topic. Chapter 3 describes our proposed method for estimating the best-time of phenological observation by information interpolation using regional amounts. Chapter 4 describes experimentally obtained results for our proposed method and a discussion of the results. Chapter 5 summarizes the contributions and future work.

2 RELATED WORK

Along with rising SNS popularity, real-time information has increased and analysis using real time data became possible. Many studies have examined efficient methods for analysis of large amounts of digital data. Some studies have been conducted to predict real world phenomena using large amounts of social big data.

Phithakkitnukoon et al. [5] analyzed the behavior of travelers such as departure place, destination, and traveling means on a personal level in detail based on massive mobile phone GPS location records. Mislove et al. [6] developed a system that infers a Twitter user's feelings from Twitter text and visualizes changes of emotion in space-time. After research to detect events such as earthquakes and typhoons, Sakaki et al. [7] proposed a method to estimate real-time events from Twitter tweets. Cheng et al. [8] estimated Twitter users' geographical positions at the time of their contributions, without the use of geotags, by devoting attention to the geographical locality of words from text information in Twitter posted articles. Although various studies have analyzed spatiotemporal data, research to estimate the viewing period using information interpolation is a new field.

3 OUR PROPOSED METHOD

This section presents a description of an analytical method for target data collection. It presents best-time estimation to obtain a guide for phenological change from Twitter in Japan. Our proposal is portrayed in Figure 1.

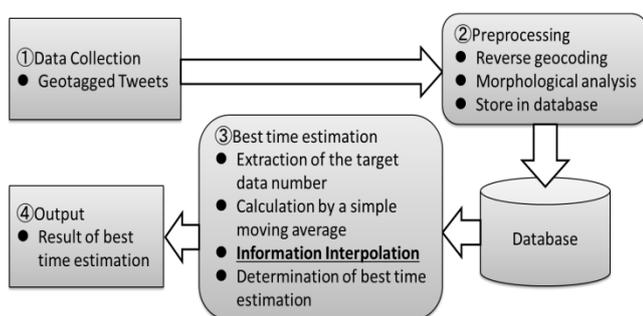


Figure 1: Our proposal summary.

We describe the best-time estimation method of organisms by analysis using moving average method for geotagged tweets containing organism names. 3.1 describe how to collect geotagged tweets to be analyzed. 3.2 describe preprocessing for conducting analysis. 3.3 describe the best-

time estimation method. In our proposed method up to now, the number of geotagged tweets was small, it was possible to estimate the best time in prefecture unit or municipality, but we could not analyze finely grain size. Therefore, by using the method using information interpolation proposed in this paper, it is possible to estimate the best time to visit by sightseeing spots with finer granularity. 3.4 describe the output of the estimation result.

3.1 Data collection

This section presents a description of the Method of (1) data collection presented in Figure 1. Geotagged tweets sent from Twitter are a collection target. The range of geotagged tweets includes the Japanese archipelago ($120.0^{\circ}\text{E} \leq \text{longitude} \leq 154.0^{\circ}\text{E}$ and $20.0^{\circ}\text{N} \leq \text{latitude} \leq 47.0^{\circ}\text{N}$) as the collection target. The collection of these data was done using a streaming API [9] provided by Twitter Inc.

Next, we describe the number of collected data. According to a report presented by Hashimoto et al. [10], among all tweets originating in Japan, about 0.18% are geotagged tweets: they are rare among all data. However, the geotagged tweets we collected are an average of 500 thousand tweets per day. As described in this paper, we use about 250 million geotagged tweets from 2015/2/17 through 2017/5/13. We calculated the best time for flower viewing, as estimated using the processing described in the following sections using these data.

3.2 Preprocessing

This section presents a description of the method of (2) preprocessing shown in Figure 1. Preprocessing includes reverse geocoding and morphological analysis, as well as database storage for data collected through the processing described in Section 3.1.

From latitude and longitude information in the individually collected tweets, reverse geocoding identified prefectures and municipalities by town name. We use a simple reverse geocoding service [11] that is available from the National Agriculture and Food Research Organization in this process: e.g., (latitude, longitude) = (35.7384446°N, 139.460910°E) by reverse geocoding becomes (Tokyo, Kodaira City, Ogawanishi-cho 2-chome).

Morphological analysis divides the collected geo-tagged tweet morphemes. We use the "Mecab" morphological analyzer [12]. By way of example, "桜は美しいです" (in English "Cherry blossoms are beautiful.") is divided into "(桜 / noun), (は / particle), (美しい / adjective), (です / auxiliary verb), (。 / symbol)".

Preprocessing accomplishes the necessary data storage for the best-time viewing, as estimated based on results of the processing of the data collection, reverse geocoding, and morphological analysis. Data used for this study were the tweet ID, tweet post time, tweet text, morphological analysis result, latitude, and longitude.

3.3 Estimating best-time viewing

This section presents a description of the method of (3) best-time estimation presented in Figure 1. Our estimation method for the best-time viewing processes the target number of extracted data and calculates a simple moving average, yielding an inference of the best time to view the flowers. The method defines a word related to the best-time viewing, estimated as the target word. The target word is a word including Chinese characters, hiragana, and katakana, which represents an organism name and seasonal change.

Next, we describe the simple moving average calculation, which uses a moving average of the standard of the best-time viewing judgment. It calculates a simple moving average on a daily basis using aggregate data by the target number of data extraction described above. Figure 2 presents an overview of the simple moving average of the number of days.

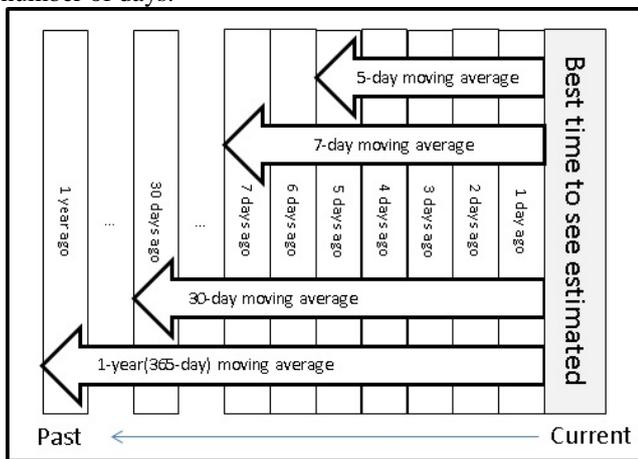


Figure 2: Number of days simple moving average.

We calculate the simple moving average in formula (1) using the number of data going back to the past from the day before the estimated date of the best-time viewing.

$$X(Y) = \frac{P_1 + P_2 + \dots + P_Y}{Y} \quad (1)$$

$X(Y)$: Y day moving average
 P_n : Number of data of n days ago
 Y: Calculation target period

The standard length of time we used for the simple moving average is a 7-day moving average and one-year moving average. A 7-day moving average is based on one week because there is a tendency for tweets to be more numerous on weekends than on weekdays. In addition, the phenological observations which are the current experiment subjects are targeting "events" that happen once a year (e.g., appreciation of cherry blossoms, viewing of autumn leaves, moon viewing). Such events are therefore based on a one-year moving average.

Next, we describe a simple moving average of the number of days specified for each organism to compare the 7-day moving average and a one-year moving average. In this study, the best time to view the period varies depending on the specified organism, the individual organism, and the number of days from the biological period.

As an example, we describe cherry blossoms. The Japan Meteorological Agency [13] carries out phenological observations of "Sakura," which yields two output items of the flowering date and the full bloom date observation target. The "Sakura flowering date" [14] is the first day on which blooming of 5–6 or more wheels of flowers occurs on a specimen tree. The "Sakura in full bloom date" is the first day on which about 80% or more of the buds are open in the specimen tree. In addition, "Sakura" is the number of days from general flowering until full bloom: about five days. Therefore, "Sakura" in this study uses a 5-day moving average, which is standard.

Next, the information interpolation method will be described. Conventionally, we estimated the best time by application of the estimation method shown in the following estimated judgment using the moving average value described above. As a result, in the analysis of a wide area such as prefecture unit, the recall rate can be estimated at about 80%. However, with an estimate of granularity such as by sightseeing spot, being unable to estimate the viewing period from the lack of data is a problem. Therefore, in this paper, we propose a method of using regional quantities that newly use information interpolation to compensate for the lack of data volume. The proposed method uses the result of reverse geocoding performed during the preprocessing in the previous section. Tweets that were judged as the same municipality by reverse geocoding are totaled for each day by city, town, or village. Then, considering the characteristic that the tweets move on a weekly basis, we obtain a 7-day moving average and set the 7-day moving average of the municipalities as the regional quantity of each region. To estimate the best time to see it, use the value obtained by adding the regional quantity of the municipality where the sightseeing spot is located to the tweet amount of the sightseeing spot to be estimated.

Next, we describe an estimated judgment of the best time for viewing, which was calculated using the simple moving average (7-day moving average, one-year moving average, and another biological moving average). It specifies the two conditions as a condition of an estimated decision for the best time for viewing. Condition 1 is the number of data one day before expression. Formula 2 is a simple moving average greater than that of the estimated best time to view date. Condition 2 is a case that follows formula 3 ((A) / (2)) or more. The short number of days by comparison of the 7-day moving average and another biological moving average is A. A long number of days is B.

$$P_1 \geq X(365) \quad (2)$$

$$X(A) \geq X(B) \quad (3)$$

Finally, an estimate is produced using conditions 1 and 2. Using the proposed method, a day satisfying both condition 1 and condition 2 is estimated as best-time viewing.

3.4 Output

This section presents a description of the method of (4) output presented in Figure 1. Output can be visualized using a best-time viewing result, as estimated by processing

explained in the previous section. A time-series graph presents the inferred results for best-time viewing. The graph presents the number of data and the date, respectively, on the vertical axis and the horizontal axis. We are striving to develop useful visualization techniques for travelers.

4 EXPERIMENTS

This chapter presents a description of the experiment to infer the best time to view flowers for the proposed method described in Chapter 3. Section 4.1 describes the dataset used for optimal time reasoning to see flowers in full bloom. As an estimation result by sightseeing spot, section 4.2 presents the estimation result without using information interpolation, and the best estimation result obtained using information interpolation in section 4.3. Section 4.4 compares the experimentally obtained results in Section 4.2 and Section 4.3.

4.1 Dataset

Datasets used for this experiment were collected using streaming API, as described for data collection in Section 3.1. Data are geotagged tweets from Japan during 2015/2/17 – 2017/5/13. The data include about 250 million items.

The estimation experiment to ascertain the best-time viewing of cherry blossoms uses the target word “cherry blossom,” which is “桜” and “さくら” and “サクラ” in Japanese. We analyzed tweets that included the target word in the tweet text.

The subject of the experiment was set as tourist spots in Tokyo. In this report, we describe “Takao Mountain,” “Showa Memorial Park,” “Shinjuku gyoen,” and “Rikugien.” Figure 3 presents the target area location. A, B, C, and D in the figure respectively denote “Takao Mountain,” “Showa Memorial Park,” “Rikugien,” and “Shinjuku Gyoen.” A and B are separated by about 16 Km straight line distance. B and C are separated by about 32 Km apart. C and D are about 6 Km apart.



Figure 3: Position of target area.

The following two experiments were conducted. The first is an experiment using the number of tweets including the target word and the sightseeing spot name without information interpolation. The second is an experiment using information interpolation. We use these datasets to estimate the optimum time for the sightseeing spots in Tokyo by experiments without information interpolation, (shown in Section 4.2) and experiments using information interpolation (shown in Section 4.3).

4.2 Estimation experiment for best-time viewing without information interpolation

In this section, we present experimentally obtained results for estimating the best time without using information interpolation from tweets containing a target word and sightseeing spot name. Figure 4 presents results for the estimated best-time viewing in 2016 using the target word ‘cherry blossoms’ in the target tourist spots. The dark gray bar in the figure represents the number of tweets. The light gray part represents best-time viewing as determined using the proposed method. In addition, the solid line shows a 5-day moving average. The dashed line shows a 7-day moving average. The dotted line shows a one-year moving average.

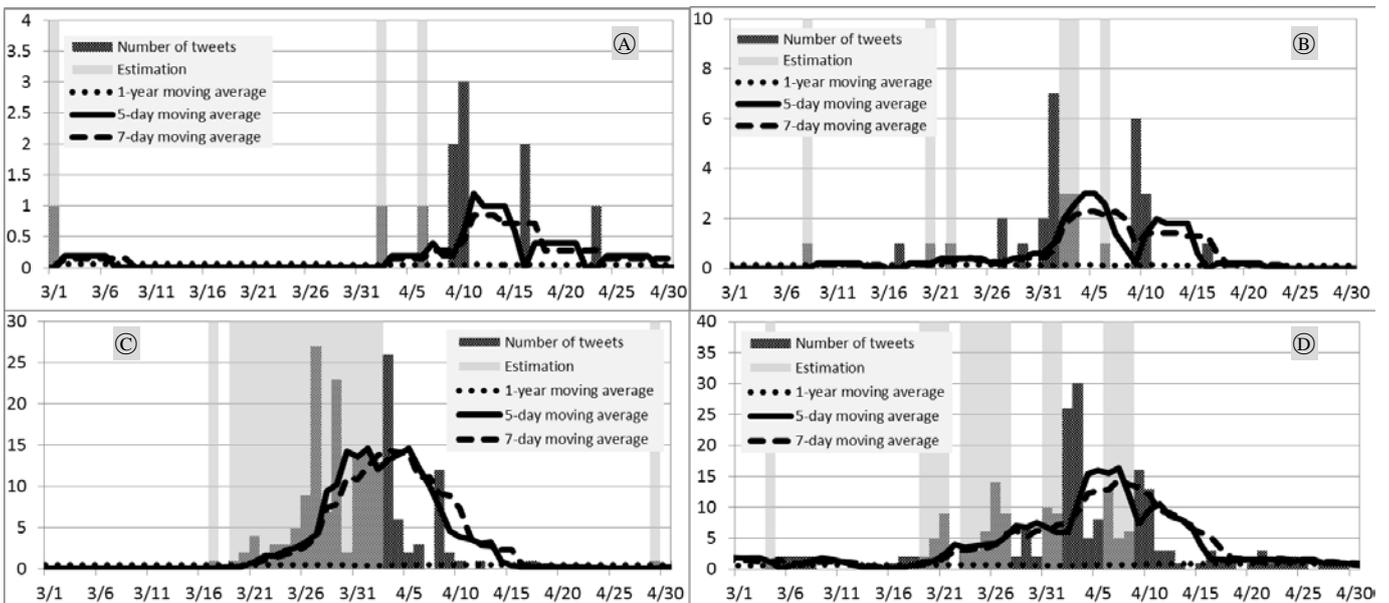


Figure 4: Experimental results obtained using tweets including the target word and the tourist spot name without interpolation.

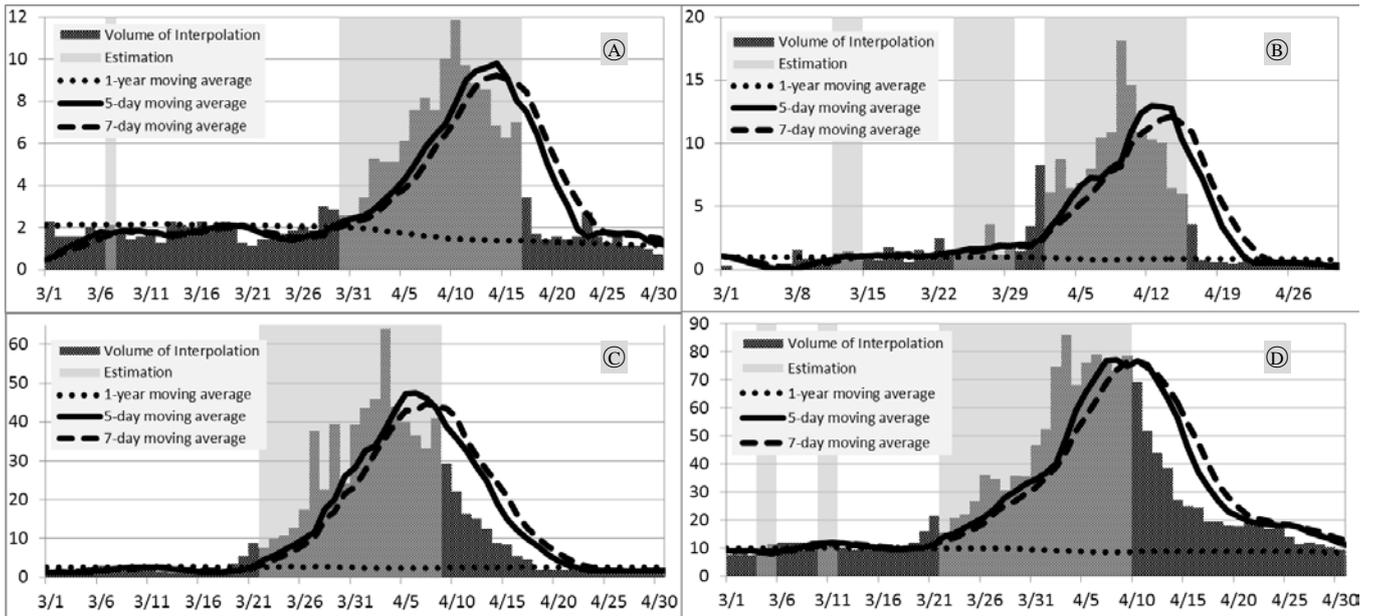


Figure 5: Experimental results obtained using tweets including the target word and the tourist spot name by interpolation.

At tourist spots targeted for the experiment in 2016, as portrayed in Figure 4, much data were obtained for C and D. The maximum number of tweets per day was about 30. These results confirmed that some estimation can be done by near-site estimation method without interpolation. However, best-time viewing cannot be done in A and B because of the very small number of tweets.

This result clarified that the method we proposed previously cannot be predictive for detailed areas such as sightseeing spots. This result is attributable to the lack of information volume.

4.3 Estimation experiment for best-time viewing with information interpolation

This section presents experimentally obtained results of estimation using information interpolation with regional quantities, which is the method proposed in this paper. Figure 5 presents results obtained using information interpolation. The notation is the same as the notation in Figure 4 in the previous section.

Apparently, A and B can produce an estimate using the proposed method by increasing the number of tweets using information interpolation with surrounding tweets. In C and D, there are days that can be determined more accurately by interpolating the number of tweets. These results demonstrate the possibility of resolving the difficulty of insufficient information when using sightseeing spot tweet data of the tourist spot area along with interpolation, and then estimating the peak period for each tourist spot.

4.4 Comparing best time for viewing estimation and observed data

Table 1 presents a comparison of experimentally obtained results of estimating the best time by information interpolation. As the table shows, Experiment 1 used the co-

occurring words in the tweet, including the sightseeing spot name coexisting with the target word "Sakura," without using the interpolation shown in 4.3. For Experiment 2, we used interpolation based on the information amount of the area including the tourist spots shown in 4.4. Numerical values in the table are the number of tweets including subject words and co-occurring words in Experiment 1. In Experiment 2, it is the sum of the number of tweets in Experiment 1 and the interpolation value of regional information amount.

Table 1: Comparison result in target areas of the best time to see the estimated and the observed data (2015)

	Takao Mountain		Showa memorial park		Rikugien		Shinjuku gyoen	
	Exp.1	Exp.2	Exp.1	Exp.2	Exp.1	Exp.2	Exp.1	Exp.2
3/18	0	2.00	0	1.00	0	1.86	2	10.57
3/19	0	2.00	0	0.57	1	3.57	2	11.86
3/20	0	1.29	1	1.57	2	5.43	5	16.00
3/21	0	1.14	0	1.14	4	8.86	9	21.43
3/22	0	1.43	1	2.43	1	7.57	0	15.43
3/23	0	1.43	0	1.43	3	10.00	3	20.43
3/24	0	1.71	0	1.57	3	10.71	3	21.86
3/25	0	1.86	0	1.43	5	12.57	6	26.57
3/26	0	2.00	0	1.71	9	17.43	14	35.86
3/27	0	1.86	2	3.57	27	37.57	9	34.71
3/28	0	3.00	0	1.14	7	22.71	2	30.57
3/29	0	2.86	1	2.14	23	39.43	7	35.57
3/30	0	2.57	0	1.43	2	24.14	2	35.29
3/31	0	2.57	2	3.43	14	39.29	10	46.57
4/1	0	3.43	7	8.29	14	43.57	9	52.14
4/2	1	5.29	3	6.14	13	45.86	26	74.71
4/3	0	5.14	3	8.71	26	64.00	30	86.00
4/4	0	5.14	0	6.43	6	44.00	5	68.00
4/5	0	6.14	0	6.86	2	40.00	8	76.00
4/6	1	7.57	1	8.00	3	36.57	13	79.00
4/7	0	8.14	0	10.43	0	33.14	5	76.86
4/8	0	7.57	0	10.86	12	41.00	6	73.14
4/9	2	10.00	6	18.14	2	29.29	16	78.43
4/10	3	11.86	3	14.57	1	22.00	13	69.29
4/11	0	9.71	0	10.86	0	16.43	3	51.57
4/12	0	8.86	0	10.29	1	15.00	3	43.71
4/13	0	8.57	0	10.00	0	12.43	1	38.29
4/14	0	6.86	0	6.43	0	8.86	0	27.00
4/15	0	6.29	0	6.00	0	8.29	1	24.86
4/16	2	7.00	1	3.57	0	5.43	3	24.43
4/17	0	3.43	0	0.71	1	4.71	2	19.14
4/18	0	1.71	0	0.57	0	1.71	2	19.14
4/19	0	1.43	0	0.57	0	1.86	0	17.86
4/20	0	1.57	0	0.43	0	1.71	1	17.57
4/21	0	1.43	0	0.57	0	1.71	3	20.00
4/22	0	1.57	0	0.57	0	1.71	1	17.71
Precision	0.51	0.77	0.57	0.74	0.95	0.84	0.80	0.82
Recall	0.06	0.58	0.22	0.44	0.39	0.58	0.56	0.83

The light gray area in the table presents the date when the satiety prediction was made using the proposed method. In addition, confirming the flowering day and full bloom period of each sightseeing spot using JMA data is difficult. Nevertheless, this experiment to evaluate SNS data for flowering is valid also for weather forecasting companies [15] and public service organizations [16] to evaluate optimum viewing times based on services and blogs that are used. Arrows indicating the flowering time can be checked manually at tourist sites. Recall and precision using the observed data and the best time to view estimated results are calculated for each target area for 2016 from 3/1 through 4/25 using formula (4) and formula (5).

$$\text{Precision} = \frac{\text{Number of days to match the observed data}}{\text{Number of days in best time to see estimated}} \quad (4)$$

$$\text{Recall} = \frac{\text{Number of days to match the observed data}}{\text{Number of days of observation data}} \quad (5)$$

Experimentally obtained results confirmed the tendency by which the relevance ratio and the recall rate became higher in Experiment 2 than in Experiment 1. In addition, A and B, which are at a higher altitude than C and D, demonstrated regional features: the best viewing time occurs later. These results confirmed the usefulness of the proposed method for best-time estimation for sightseeing spots using information interpolation along with regional data.

5 CONCLUSION

As described herein, to improve best-time estimation accuracy and thereby enhance tourist information related to phenological observation, we proposed an information interpolation method. The proposed method showed optimal time to see flowers at sightseeing spots by interpolating information using the 7-day moving average of the number of tweets of municipalities, including those of sightseeing spots. This method can estimate the best time for sightseeing spots with fine granularity, giving predictions in units required for sightseeing.

The results of cherry blossoms experiments conducted for tourist spots in Tokyo in 2016 using the proposed method confirmed improvement of the estimation accuracy when using information interpolation. The proposed method using information interpolation for tweets related to named organisms (sakura trees) might improve the real-world accuracy of estimating best times. We confirmed the possibility of applying this proposed method to estimation of viewpoints and lines of sight in areas and sightseeing spots with few tweets and little location information.

Although the proposed method succeeded in interpolation of information and highly accurate estimation, it may be due to the fact of cherry blossoms blooming with short-term changes and the fact that the public interest is high. Therefore, it is necessary to verify whether the same results can be obtained by another biological season observation as a future task. Future studies must also examine automatic extraction of target words and a method to perform future predictions in real time. Eventually, this system might be extended to a system by which travelers can obtain travel-

destination-related event information and disaster information in real time.

ACKNOWLEDGEMENTS

This work was supported by JSPS KAKENHI Grant Number 16K00157, 16K16158, and a Tokyo Metropolitan University Grant-in-Aid for Research on Priority Areas “Research on Social Big Data.”

REFERENCES

- [1] Japan Tourism Agency, Tourism Nation Promotion Basic Law, URL <<http://www.mlit.go.jp/kankocho/en/kankorikkoku/index.html>> (2007).
- [2] Ministry of Economy, Trade and Industry, Study of landing type IT tourism and attract customers service, URL <<http://www.meti.go.jp/report/downloadfiles/g70629a01j.pdf>> (2007) (in Japanese).
- [3] Twitter, URL <<https://Twitter.com/>> (2014).
- [4] M. Endo, Y. Shoji, M. Hirota, S. Ohno, and H. Ishikawa, On best time estimation method for phenological observations using geotagged tweets, IWIN2016 (2016).
- [5] S. Phithakkitnukoon, T. Teerayut Horanont, A. Witayangkurn, R. Siri, Y. Sekimoto, and R. Shibasaki, Understanding tourist behavior using large-scale mobile sensing approach: A case study of mobile phone users in Japan, Pervasive and Mobile Computing (2014).
- [6] A. Mislove, S. Lehmann, Y.Y. Shn, J.P. Onnela, and Rosenquist, Understanding the Demographics of Twitter Users, Proceeding, Fifth International AAI Conference on Weblogs and Social Media (ICWSM'11), pp. 133-140 (2011).
- [7] T. Sakaki, M. Okazaki, and Y. Matsuo, Earthquake shakes Twitter users: real-time event detection by social sensors, WWW 2010, pp. 851-860 (2010).
- [8] Z. Cheng, J. Caverlee, and K. Lee, You are where you tweet: a content-based approach to geo-locating twitter users, Proceedings of the 19th ACM International Conference on Information and Knowledge Management (2010).
- [9] Twitter Developers, Twitter Developer official site, URL <<https://dev.twitter.com/>> (2014).
- [10] Y. Hashimoto and M. Oka, Statistics of Geo-Tagged Tweets in Urban Areas (<Special Issue>Synthesis and Analysis of Massive Data Flow), JSAI, Vol. 27, No. 4, pp. 424-431 (2012) (in Japanese).
- [11] National Agriculture and Food Research Organization, Simple reverse geocoding service, URL <<http://www.finds.jp/wsdocs/rgeocode/index.html.ja>> (2014).
- [12] MeCab, Yet Another Part-of-Speech and Morphological Analyzer, URL <<http://mecab.googlecode.com/svn/trunk/mecab/doc/index.html>> (2012).

- [13] Japan Meteorological Agency, Disaster prevention information XML format information page, URL <<http://xml.kishou.go.jp/>> (2011).
- [14] Japan Meteorological Agency, Observation of Sakura, URL <<http://www.data.jma.go.jp/sakura/data/sakura2012.pdf>> (2016).
- [15] Weathernews Inc., Sakura information, URL <<http://weathernews.jp/sakura>> (2016).
- [16] Japan Travel and Tourism Association, Whole country cherry trees, URL <<http://sakura.nihon-kankou.or.jp>> (2015).

Automatic Recording of Actors' Body Expressions with Action Cueing

Takayoshi Takano^{*}, Hiroshi Shigeno^{**}, and Ken-ichi Okada^{**}

^{*}Graduate School of Science and Technology, Keio University, Japan

^{**}Faculty of Science and Technology, Keio University, Japan
{takano, shigeno, okada}@mos.ics.keio.ac.jp

Abstract—The body expression of an actor is one of the important factors in theatrical performance. The main purpose of self-practice of actors is to check and review such important factors. Therefore, it is important to keep a record of practice to improve the quality of a theater. Currently, the recording of practice is done mainly in the form of writing to the script, but it is difficult to record the physical expressions of an actor accurately. We focused on motion capture as a way to record body expressions to deal with this difficulty. In this paper, we constructed a system aimed at automatic recording of actors' body expression in theatrical practice by action cueing. By notifying the turn of acting for an actor using smart watch, our system can record the physical expression while supporting the actors' practice. We conducted an experiment to evaluate the efficiency of this system. From the results, it was confirmed that this system can record the body expression with a small burden.

Keywords: motion capture, theatrical performance, body expression, smart watch, action cueing.

1 INTRODUCTION

Theater organizations practice many times to make the performances better. In practice, a director instructs actors about important elements in performance. On the other hand, the actor reads into the intention of the director and tries to express it with the body movements. According to Goan et al., "Timing of each action", "Body expressions", and "Feelings" are important [1]. A director mainly instructs actors about these important elements during practice.

- Timing of each actions

Timing of the start of a line creates a span between lines. Different from a picture work such as movies, spans are transmitted to the audience in real time. Spans are the time the audience uses imagination about the kind of relationship between characters. Therefore, spans are very important element being concerned with success or failure of performances. In actual practice, instructions from a director to an actor about spans are the most frequent. A director coaches actors about it very finely. As an example, "a little early", "immediately begin to say when the line of that person is finished", "open for two seconds." are given. By taking the appropriate spans, performances become better, and it is possible to avoid lines become monotonous.

- Body expressions

The actors express the feelings of the roles playing with their own bodies. Frequently, the director instructs the actor

how to move in order to express the feelings of the role. Also, it is a very important factor how the actors look like from the audience. Therefore, actors often check their movements with a mirror.

- Feelings

Feelings in performances are expressions and perceptions that actors feel. An actor must speak lines with getting into the character. When the remark of actors is not enough of feelings, it is given a worthless performance. Therefore, directors coach for actors about feelings frequently. The grounds of remark of actors are feelings that the character has.

Many theatrical organizations have problems of lack of practice places and opportunities [2], therefore they encourage actors' self-practicing to improve the quality of performances. In self-practice, actors mainly review the previous lesson and prepare for the roles by reading the script for the next lesson. Therefore, keeping a record of practice is very important to improve the quality of the performance. The following methods can be considered for recording performance.

- Writing to the script by text

Recording is done by handwriting. Currently, this is the most common recording method. This method is difficult to record nonverbal elements such as body expression. In addition, there is a disadvantage that an actor must temporarily stop the performance to record.

- Movie

In this method, it is possible to record a lot of information and record nonverbal elements. Indeed, there are many actors to use for review by taking movies of their own acting. However, the recorded movie has no use other than watching it. In addition, there is a disadvantage that confirmation becomes complicated in proportion to the recording time. For example, if you record a two hour lesson by a video, it is hard work to find a few seconds of acting that you would like to check.

- Motion capture

This method is suitable for recording non-verbal elements such as body expressions. Also, recorded data can be applied to various things. For example, it is possible to visualize a recorded motion data and confirm the performance from an arbitrary viewpoint. In addition, if you record the performance separately for each actor, it is possible to easily search for performance that the actor

wants to confirm. On the other hand, there is a disadvantage that cost of motion capture equipment is high, and in some cases, additional time is required for motion capture.

Taking into account the above, we focused on motion capture because the record of the physical expression of the actor can be handled as digital data. Considering the use of the record of physical expression for practice, we think it is important that the data format should be highly versatile. In the following chapter, we introduce Digital-Script that handles elements necessary for theater as digital data.

In this paper, we construct a system aimed at automatic recording of actors' body expressions in theatrical practice while supporting actors' learning. In recording practice, it is difficult to grasp the timings of actors' performances. To deal with this problem, we utilize action cues for supporting theatrical practice. We assume that action cues are not only useful for supporting theatrical practice but also for grasp timings of actors' performances. By automating recording body expressions of actors', it is unnecessary to devote time for motion capture separately from the practice time, and it is possible to reduce the workload of recording body expressions.

In order to evaluate user's workload of this system, we conduct experiments in two situations: recording by this system and by a human-shaped input device. In both situations, the subjects participate as the director and record the physical expression of actors. Two experimenters participate as actors. In the experiments, we compared workload and time required to record actors' body expressions. From the results of the experiments, we confirmed that this system can record body expressions by lower workload and shorter time. The rest of this paper is organized follows: Section 2 introduces Digital-Script, and Section 3 describes related work. Then, the detail of our system is described in Section 4, Section 5 shows evaluation of this system. Finally, we conclude in Section 6.

2 DIGITAL-SCRIPT

We propose Digital-Script which is added with the data necessary to performances. Digital-Script is handled as a database, and is different from the conventional scripts. Theatrical production has technical elements and actor's performance elements. Technical factors are such as sound effect, lighting effect, stage carpenter and stage design. In actor's performance, "Timing of each actions", "Body Expressions", and "Feelings" are very important elements. However, as mentioned above, it is difficult to read these factors from a script written only in text. Therefore, in order to visualize these important elements, we try to handle them as digital data. Upon making database of scripts, we compose Digital-Script by text and numerical value in order to make the system which has a variety methods of the visualization to each needs, such as voluntary exercise of actors, directions input of a director, and stage confirmation. In addition, in order to help actors grasp the timing of each action that changes on the stage, we add time information to each data.

Digital-Script is handled as database which is implemented by MySQL. Figure 1 shows the configuration of the Digital-Script database. Digital-Script is distributed by scene unit. Scene data manages a certain scene and includes "SceneID", "Title", "SceneName", and Actor data. In the following, we explain Actor data in detail. In Actor data, the peculiar ID and name are assigned to every character. In addition, "SAY", "FEELING", "MOVE", and "SEE" are included in it.

- SAY
SAY data are about actors' speech. This means that an actor speaks A when the elapsed time gets t sec. Strength means that an actor should strongly/weakly read an appointed part of lines A. Long means that an actor say lines A in l second.
- FEELING
FEELING is a data about the feelings of an actor. This means that an actor has feelings E at the time of t second.
- MOVE
MOVE data are about actor's standing position. This means actor's movement to (x, y, z) coordinates during elapsed time between t1 sec to t2 sec.
- SEE
SEE data are about actor's head direction where the actor should look at. This means that the actor looks at the designated direction when the elapsed time get t sec. The direction where the actor should look at is designated by (x, y, z) coordinates.

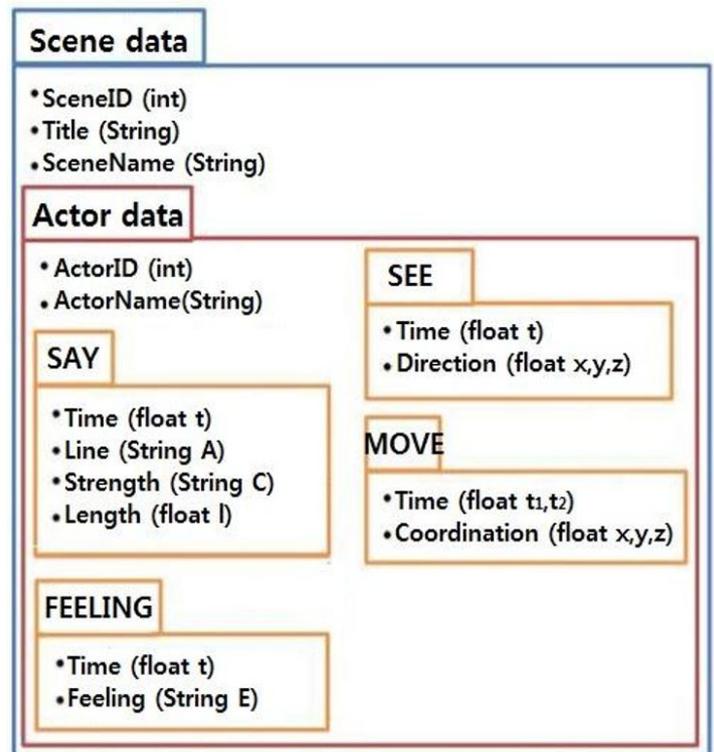


Figure1: Components of Digital-Script Database

3 RELATED WORK

In this chapter, we introduce several studies to support theatrical activities.

3.1 Recording of dancing movements

Hachimura et al. tried to motion capture the behavior of Japanese dance, which is a traditional Japanese art, to make it useful for inheritance and preservation [3]. In addition, Hachimura and colleagues analyze the characteristics of "Okuri" which is the basic behavior of dancing by using motion-captured data. There are also many studies using motion capture in the field of ballet and dance. Saga and colleagues have studied using motion capture for simulation of ballet choreography [4]. Their study's goal is to develop an automatic composing system for ballet by using motion Data archives. However, these studies use large-scale equipment for motion capture, and Hachimura also states that the operation of these equipments requires experts with special skills.

3.2 Action cueing

Takatsu et al. applied smart watch to support the practice of dramas [5]. In order to practice smoothly, it is important for the actors to learn the correct order of performance. However, in the early steps of practice, it is difficult for actors because a play script has a lot of information to learn. In this research, Smart watch has the role of telling the information to the actors and informing them in the correct order. Smart watch is a way to convey information without putting a physical burden on the actor. This research makes it easier for actors to perform in the proper order, and actors can concentrate more on acting.

3.3 Digital-Script

We introduce two studies to support practice of theater by using Digital-Script.

Shimada et al. proposes the voluntary practice system [6] reflected a direction of a director. This study focuses on timing of an action such as beginning of lines and movements. This system visualizes Digital-Script in two ways like figure 2. In this display, timing of each action is showed by a time-line and spatial information is showed by 3DCG. By using this system, it is possible for an actor to practice while grasping progress of a performance.

Theater organizations may not practice anytime with all members because they don't have their own practice space and most actors have another job. By using Digital-Script, theatrical practice support system [7] enables actors to practice performance in the situation that a director or a part of actors cannot participate. This system needs monocular Head Mounted Display (HMD). In this HMD, virtual actor that plays absent actor's role is displayed, and a user is coached automatically by detecting and comparing actor's movement with information contained in Digital-Script like figure 3.

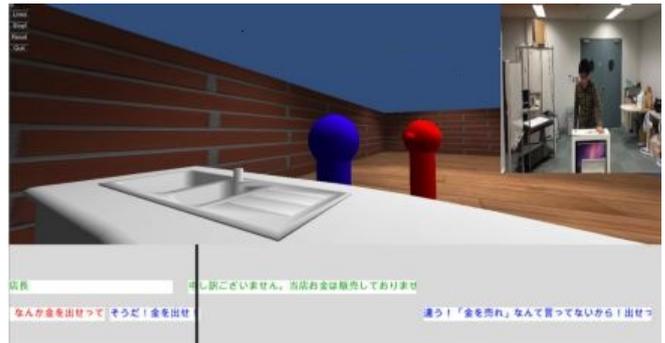


Figure 2: Voluntary practice system [7]

4 AUTOMATIC RECORDING OF ACTORS' BODY EXPRESSIONS WITH ACTION CUEING

Recording performance is important in the practice of dramas. However, with the current method, in practice of acting, recording body expressions is not easy. Therefore, we propose the system that can automatically record actors' body expressions while supporting for actors' learning. This system can notify action cueing to actors during recording body expressions. The contributions of this research are as followings.

- Automation of recording body expression
By automation of the recording of the actors' physical expressions, there is no need to perform additional work for recording. This system does not require special skills or more time than practice time. This contribution can solve the problem that the recording operation shortens the practice time.

- Digital-Script
Scripts are what represent story of the theater. However, since scripts are written in text, actors cannot intuitively grasp the non-verbal elements such as the movement of the body and the standing positions of the actors. Therefore, we propose the motion capture system for the physical expressions of actors. We treat this recorded motion data as Digital-Script and save it in a form that can be used for various purposes.

- Notifying of the order of performance of actors
In this system, smart watch notifies the order of performance of each actor individually. Smart watch shows what the actor will do next and the next line. Smart watch's vibration informs the order of acting, so the actor does not need to see the screen. By using this system in the early stage of practice, the actor can learn the flow of acting by notifying action order. This system can automatically record the body expression while actors concentrate on performance.

4.1 System Overview

This system automatically records the performance without a burden for an actor or a director. Also, this system notifies actors of the order of acting and supports the practice of

actors. Figure 3 shows the usage of this system. This system requires PC, Kinect, tablets, and smart watches.



Figure 3: Usage image of this system

The director uses the web application on the PC. The figure 4 shows the screen of the web application. On this screen, the result of reference to Digital-Script database is displayed. Lines, timings, actor's names, actions are shown. The director checks the performance of the actor while confirming the flow of the script on this screen. By pressing the start button, practice begins along the data of Digital-Script. The director can start practicing from any line of speech. Also, the director can delete or add lines on this application.

EDIT		SCRIPT				
Number to insert: 2		Order	Timing	Actor	Line	Movement
Contents: 指示を出しながらボールを insert		1	0.0	B	右！右！ああ！もっと左！	初期位置
		2	4.0	A	ダメだ！見えない！まーっとく見えない！	0
		3	8.0	B	見えるって！目を細めたら見えるって！	0
		4	12.0	A	モザイクじゃないんだから！そんなので見えないよ！	0
		5	16.0	B	だいたいさ、見えない見えないって、目の前にいる加藤さんに失礼だろ？	0
		6	22.0	A	え！？あ！すいません！いたんですか！	0

Figure 4: Web application

When the performance starts, this system tracks progress of performance automatically according to Digital-Script, and notification is sent to the actor's smart watch. At that time, all movements of actors are motion captured by Kinect and saved as Digital-Script. On the other hand, the actor wears a smart watch and practices. The smart watch notifies by vibrating according to the behavior of each actor and the timing of speech. This timing of notification is based on the Digital-Script database. Figure 5 shows the notification display of smart watch.



Figure 5: Notification window of smart watch

4.2 Use case

We assume that this system is used in the following cases.

- Early stages of practice
This system is supposed to be used at the stage of practice. In early stage of practice, actors do not remember the order of acting perfectly. Normally, at this stage, the actors practice while having a script, and receive instruction from the director.
- 4 meters square space
This is the extent of space where actors can practice while moving. This range of space depends on specifications of Kinect.
- Participating of the director in practice
We assume that the director participates in the practice and instructs the actors' acting. The director has the right to decide all performance. Therefore, we assume that theatrical organizations use this system with the aim of recording of performance.

4.3 Implementation

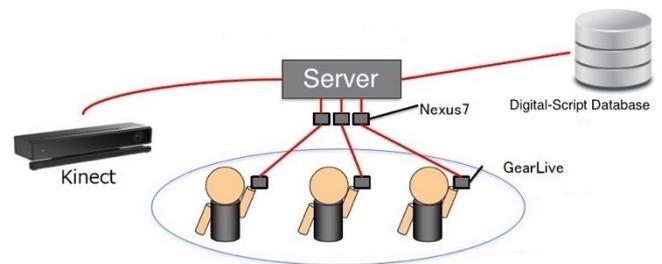


Figure 6: Hardware architecture

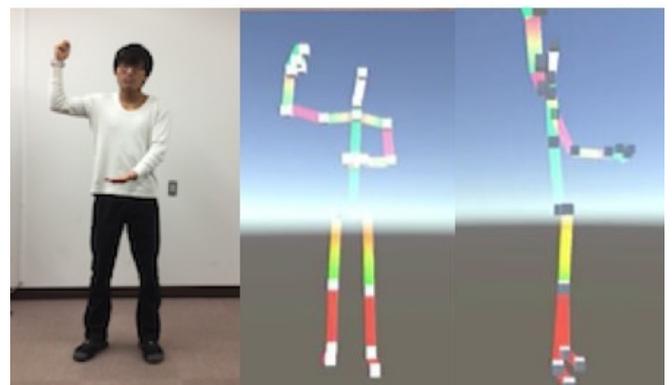


Figure 7: Visualized image

Figure 6 shows the hardware architecture of this system. This system requires PC, Kinect, tablets, smart watches. We used Samsung Gear Live for smart watches and Nexus 7 for Tablets. Samsung Gear Live is connected to the PC server by WebSocket. In order to connect quickly and support multiple devices, we use WebSocket. WebSocket can keep socket open with bidirectional connection and support stable connection in multi device environment. This

was implemented by node.js and was connected to the device using GitHub and heroku.

In this system, each actor wears this smart watch and receives action cueing by vibration. Samsung Gear Live is paired with a Nexus 7 tablet. Since only one smart watch can be controlled with one tablet, we used multiple tablets connected by Bluetooth. We use Android Wear 5.0.1 as the implementation platform and the Android Java API v21.1.2. The processor is 1.2 GHz Qualcomm Snapdragon 400, and the display is Super AMOLED 1.63 type. Tablet OS is Android Wear 4.4.3, 1.3 GHz NVIDIA Tegra 3 mobile processor. Kinect for Windows API v2.0 was used as a depth camera to recognize the movement of each actor. Kinect has a depth camera, RGB camera, multiple eye ray microphone, and processor. Kinect can recognize the gesture, the position, and height of the users. Kinect can detect up to six people simultaneously. During practice, Kinect will always detect the actor and start recording at the timing when the action cue is notified to the actor. The recorded skeleton coordinates are output as a CSV file and finally converted to a Digital-Script format. Figure 7 shows an example of visualized Digital-Script data.

5 EVALUATION

In this chapter, we describe experiment to evaluate workload of this system in recording actor's body expressions.

5.1 Subjects

10 university and graduate school students participated in this experiment. One subject participated in one experiment.

5.2 Method

As a director, the subject instructed the actors about performances while watching the script. Two experimenters participated in the experiments as the actors. Subjects recorded performances of actors. We conducted the experiment in two situations: Experiment A and B. In Experiment A, subjects recorded body expressions of actors with this system, and in Experiment B, subjects record it with human-shaped device. Five of the ten subjects belonged to Group 1. Firstly, Group 1 participated in Experiment A and then did Experiment B. The remaining 5 subjects belonged to group 2. Group 2 first participated in Experiment B and then did Experiment A.

In Experiment A, the subject participated in this Experiment As a director. First, the subject saw the performance of the actor. Subsequently, the subject instructed the acting of the actors. The content of the instruction was what the experimenters informed the subject in advance. After those instructions, the actors performed acting according to instructions. The subject recorded their performance using this system. We measured the time required to record performance. To summarize, Experiment A proceeded in the following procedure.

1. Explanation of the flow of the Experiment and the operation of this system

2. Watching practice
3. Instructing to the actors
4. Recording practice

In Experiment B, like Experiment A, the subject participated in this experiment as a director. The subsequent flow is the same as Experiment A, but in this experiment this system was not used. First, the subject saw the performance of the actor. Subsequently, the subject instructed the acting of the actors. After those instructions, the actors performed acting according to instructions. We recorded that situation in video. The subject created motion data using QUMARION (see figure 8) while referring the video. QUMARION is a human type input device developed and sold by Celsys [8], Inc. It is a device used for creating motion and pose of 3D animation. The size of this device is $29 \times 30.8 \times 10.8$ (cm) and the weight is 255g. It has 16 joints and 32 sensors and can create motion intuitively and easily. We measured the time required by subjects to make motion using this device. To summarize, Experiment B proceeded in the following procedure.

1. Explanation of the flow of the experiment and the operation of this system
2. Watching practice
3. Instructing to the actors
4. Practice with recording by video
5. Creating motion data by QUMARION with reference to movie



Figure 8: QUMARION [8]

We used a script for each experiment. We used two scripts, X and Y, from free script download service [9]. Details of the two scripts are shown in the table 1.

Table 1: Details of scripts

	Length	Number of lines	Number of tasks
Script X	150 sec	30	6
Script Y	150 sec	32	6

5.3 Evaluation Items

In both experiments, we evaluated the workload and required time in the recording of the physical expression of the actors.

To evaluate workload of with this system, we used NASA - Task Load Index (NASA-TLX, here in after referred to as TLX). TLX is one of the most frequently used subjective

workload metrics. TLX has 6 scales: Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, and Frustration. They are rated for each task within a 100-points range with 5-point steps. Figure 10 shows part of the answer sheet for these scales. Subjects fill in this line the degree of involvement in the workload for each scale. We measure the position of the mark on the line to get a rating of 1 to 100. We calculated the average using WWL (weighted workload). A small WWL value means that the participant's workload is small.

In Experiment A, we measured the time taken for recording of the physical expressions of the actors. On the other hand, in Experiment B, we measured the time taken to create motion data by QUMARION.

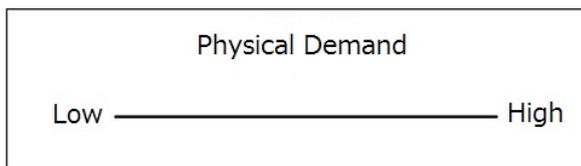


Figure 9: Workload metrics

5.4 Results and Discussion

Table 2 shows the results for group 1 workload, and Table 3 shows group 2 results. Table 4 summarizes the results on the time required for each group of experiments.

Table 2: Workload of group 1

Experiment	Script	WWL(Weighted workload)
A	X	33.8
B	Y	56.5

Table 3: Workload of group 2

Experiment	Script	WWL(Weighted workload)
A	Y	38.3
B	X	51.1

Table 4: Results of time required

Experiment	Group 1 (sec)	Group 2 (sec)	Ave. (sec)
A	223.4	195.2	209.3
B	827.2	626.0	726.6

From table 2 and 3, it can be seen that the values of WWL are smaller in Experiment A in both groups. From this, we found that this system can record actors' body expressions with low workload.

From table 5, we see that the time taken by Experiment A is shorter. From this result, it can be said that subjects record body expressions with this system by the shorter time. We assume that one of the reasons for reducing workload is shorter time required.

From table 1 and 4, we found that Experiment A takes about 209 seconds compared with the time of the scripts of 150 seconds. The time required by this system is slightly longer than the script time, but we think that recording can be completed in a sufficiently short time.

6 CONCLUSION

Presently, recording practice is mainly done by writing to the script by text. However, text is not suitable for recording the physical expression of the actor, and it is difficult for the actors to review during the self-practicing. So, we tried using motion capture. Generally, motion capture requires large equipment and load. In this thesis, we propose a system that can automatically record body expressions during practice by motion capture. In addition, by notifying the actor of the action cue by using the wearable device, this system also supports the practice of the actor.

We conducted an experiment to evaluate the workload of this system. As a result, we confirmed the tendency for this system to record the motion data of actors in a short working time and low workload.

REFERENCES

- [1] Goan, M. System dynamics on the creation of dramamaking processes. *Jpn. Cogn. Sci. Soc. Cogn. Sci.* 14(4), pp. 509–531 (2007). (in Japanese)
- [2] Agency for Cultural Affairs, Survey on training and utilization of human resources related to demonstration artists. (in Japanese)
http://www.bunka.go.jp/seisaku/bunkashingikai/seisaku/07/03/pdf/shiryo_3.pdf
- [3] K. Hachimura, Digital Archiving of Dancing by Using Motion Capture, Computer Vision and Image Media, pp. 1-8, (2007). (in Japanese)
- [4] A. Soga, B. Umio, Creative Simulation by Analyzing Choreographic Elements of Classical Ballet, *IPSJ SIG Computers and the Humanities*, pp. 45-52, (2008). (in Japanese)
- [5] R. Takatsu, N. Katayama, H. Shigeno, and K. Okada. A Wearable Action Cueing System for Theatrical Performance Practice, *Collaboration Technologies and Social Computing: 8th International Conference, CollabTech 2016, Proceedings*. Vol. 647. Springer, (2016).
- [6] Shimada, M., Fujishige, S., Okada, K.-I., Supporting actor's voluntary training considering the direction of drama. *IPSJ Trans. Digit. Content* 4(1), pp. 1–9 (2016). (in Japanese)
- [7] Shimada, M., Takayoshi T., Hiroshi, S., Okada, K.-I., Supporting Theatrical Performance Practice by Collaborating Real and Virtual Space. *Collaboration Technologies and Social Computing: 8th International Conference, CollabTech 2016, Proceedings*. Vol. 647. Springer, (2016).
- [8] CELSYS, Inc.
<https://www.celsys.co.jp/>
- [9] Free script download service: Hariko no tora no ana.
<https://haritora.net/script.cgi>

Discovering Hotspots Using Photographic Orientation and Angle of View from Social Media Site

Masaharu Hirota[†], Masaki Endo[‡], Daiju Kato[#], and Hiroshi Ishikawa^{*}

[†]Faculty of Informatics, Okayama University of Science, Japan

[‡]Division of Core Manufacturing, Polytechnic University, Japan

[#] WingArc 1st Inc., Japan

^{*} Faculty of System Design, Tokyo Metropolitan University, Japan

hirota@mis.ous.ac.jp, endou@uitech.ac.jp, kato.d@wingarc.com, ishikawa-hiroshi@tmu.ac.jp

Abstract - Hotspot is interesting places for many people to do sightseeing. Visualization of hotspots reveals user interests, which is important for industries such as tourism and marketing research. Hotspot is classifiable to two types: area of interests and shooting spot. This paper introduces a new method for extracting the area of interests using various metadata annotated with a photograph from photo-sharing sites. Although several social-based techniques for extracting hotspots have been proposed using photographic location, those most methods cluster photographs based solely on the density of geographic proximity. However, in almost cases, a hotspot and shooting point where photographs were taken are distant. Also, when a photograph was taken, an angle of view, which the breadth of a subject as seen by a camera system, was measured by a camera. Therefore, we propose to extract the area of interests using photographic orientation and angle of view in geo-referenced and oriented photographs. We demonstrate our approach by extracting the area of interests using photographs annotated with those metadata from Flickr.

Keywords: area of interest, shooting spot, photograph location, photograph orientation

1 Introduction

According to the increasing popularity of mobile devices such as digital cameras and smartphones, numerous photographs taken by photographers have been uploaded to photo-sharing web services such as Flickr [1] and Panoramio [3]. Recently those devices have included embedded global positioning system (GPS). Using them, photographers can readily take photographs with photographic metadata. A photographic location which is one of a type of the metadata observed by GPS shows a place at which photographer took a photograph. Also, photographic orientation indicates the direction in which the photograph was taken from the photographic location. Particularly, photographs with a photograph-orientation feature have become numerous recently. Also, many photographs on photo-sharing sites have metadata that are annotated by users through social tagging.

Many people might take photographs of subjects or landscapes satisfying their own interests. Then, they upload those photographs to the sites. As places at which many photographs have been taken, these locations might also be interesting places for many people to visit. In this paper, we define such places as a hotspot. Also, we can extract the interesting places

using photographs obtained from photo-sharing sites. Therefore, some methods have been proposed to extract hotspots from photographs at photo-sharing sites [14], [25], [27], [28], [24]. The extracted hotspots might reflect people's interests, or be useful for marketing research, spatial analysis, and so on [23], [11]. Analyzing such areas is necessary for industries such as those related to tourism [23],[21]. Furthermore, tourist attraction recommendation systems such as [16], [21] can use this approach. By presenting hotspots to people who visit a city for the first time, our approach assists tourism.

A hotspot is classifiable into two types: area of interest (a place captured in the photographs), or shooting spot (a place to take photographs). Figure 1 shows examples of shooting spots and area of interests. In figure 1, the blue circle shows photographic location, and break line shows the photographic orientation and angle of view. Also, the cell filled by green shows shooting spots, and the cell filled with orange shows an area of interest. We define a shooting spot as one type of hotspot in an area where many photographs have been taken. When a photographer takes a photograph of attractive spots such as landmarks or landscapes, they take the photographs at a place which is distant from the spots. Also, an area of interest (or point of interest) of the other type of a hotspot is attractive as tourist spots for many people (e.g., Colosseum, Statue of Liberty). In such areas, many photographers take a lot of photographs inside or nearby from distant places. Such places are also extracted as hotspots and are defined as an area of interest. In addition, hotspots occur because of an event that might occur. When an event such as reworks display happens, many people take photographs related to the event.

Although many methods to extract and visualize the area of interest from photo-sharing sites have been proposed, those studies extract area of interest using photographic location where photographs were taken. Almost researchers to discover area of interests extract places where the density of photographic location is higher than other places using the density-based algorithm like DBSCAN [9] and P-DBSCAN [14]. However, in almost case, a photographic subject of a photograph is several meters away from a place where a photographer took the photo, like Figure 1. As a result, a place extracted based on the density of photographic location is not the area of interest, but shooting spot. Therefore, we extract area of interests using photographic orientation and angle of view of a photograph. Modern cameras and smart-

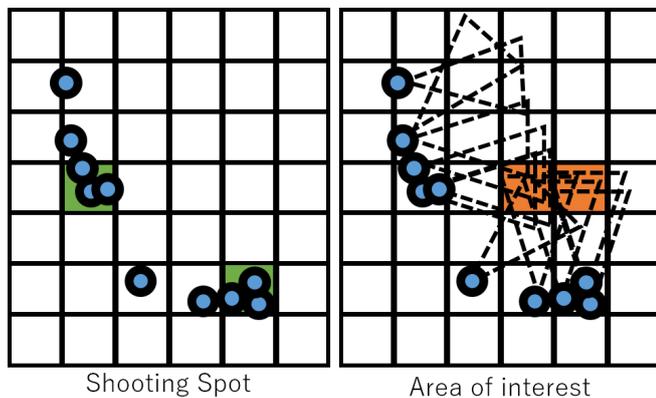


Figure 1: Area of interest and Shooting spot

phones equipped with a digital compass can add the metadata of photographic orientation taken by a user. Photographic orientation presents the direction in which the photograph was taken. Also, the angle of view is the angular degree of a given scene that is included in the photograph taken with a camera. In this paper, we propose a novel clustering method to extract area of interests using photographic orientation and angle of view. In our approach, we use pseudo triangle which is created by the photographic orientation and angle of view. We classify the area into an area of interest or not based on the degree of overlap of the triangles.

The remainder of the paper is organized as follows. Section 2 presents works related to extracting hotspots and area of interests. Section 3 presents a description of our proposed method to extract area of interests using photographic orientation and angle of view. Section 4 explains several examples of our proposed system and presents a discussion of the results. Section 5 concludes the paper with a discussion of results and future works.

2 Related Work

2.1 Clustering algorithms to extract hotspots

Some methods have been proposed to extract hotspots from the many photographs with the photographic location which are available on photo sharing sites. There are two main approaches to extract hotspots using the photographic location.

First is the density-based clustering such as DBSCAN [9] or meanshift [7] to extract hotspots from a dataset which includes huge photographs annotated with photographic location. Crandall et al. presented a method to extract hotspots using meanshift based on many photographs annotated with photographic location from the photo-sharing site [8]. Kisilevich et al. proposed P-DBSCAN, which a density-based clustering method improved DBSCAN for the definition of reachable point, to extract hotspots using the density of photographic locations [14]. Ankerst et al. proposed a clustering method of OPTICS which is a variation of DBSCAN to create a cluster using different subspaces extracted from various parameter [5]. Sander et al. proposed GDBSCAN of generalized DBSCAN which extends to enable the corresponding to both spatial and their non-spatial features [22]. Shi et al. proposed density-based clustering method to extract places of interest using spatial information and the social relationships between users [24].

The other approach to extract hotspots is grid-based clustering algorithm. Grid-based clustering is that the data space is quantized into a finite number of cells which is formed by the grid structure. Also, in the grid-based clustering, identifying which a cell is a cluster or not is the number of data included in the cell. The main advantage of the algorithm of grid-based clustering is a fast processing time, which most of the algorithms achieve a time complexity $O(n)$ of where n is the number of data (e.g. DBSCAN is $O(n^2)$ using kd-tree) [17]. Also, the performance of clustering depends only on the size of the grids which is usually much less than the data objects [19]. Additionally, almost grid-based clustering algorithms are easy to extend parallelization, because each cells are independent when the algorithm detects whether the cell is defined as a cluster or not. Agrawal et al. proposed CLIQUE to extract clusters within subspaces of the dataset using apriori-like technique [4]. Wang et al. present STING which represents a grid-based and density-based approach [29]. Chang et al. proposed Axis-Shifted Grid-Clustering algorithm, which is a dynamic adjustment of the size of the original cells in the grid and reduction of the weakness of borders of cell [6].

Our proposed method to extract area of interests calculates the density of the area which photographers took photographs. We extract clusters as hotspots based on density in a grid space. To calculate a density of each cells in the grid, we use overlaps of the area taken by photographers, calculated by photographic location, orientation, and angle of view of photographs. In this paper, our proposed method adopts the number of overlaps of pseudo triangles at which photographs were taken in each cells to identify a cell is area of interests or not. This reason is that our approach uses the pseudo triangles to extract area of interest. The location is a point, but calculated pseudo triangles present an area. Therefore, the amount of data applied to clustering algorithm is more huge, comparing to the approach using photographic location. Therefore, the time complexity of the advantage of the grid-based approach is important to treat huge photographs.

2.2 Extracting hotspots based on photograph orientation

According to photographs with a photographic orientation have become more commonly available, photographic orientation is used for extracting hotspots. Lacerda et al. proposed a method for extracting hotspots using photograph orientations [15]. This method calculates the intersections between lines of photographic orientation of many photographs. The intersections are clustered using DBSCAN. In addition, Thomee et al. proposed a method for consideration of inaccuracies affecting GPS location measurements [27].

However, those methods to extract hotspots using photographic orientation do not consider the angle of view of photographs. The angle of view is important information which shows an area of user interest. Therefore, we propose a method to extract area of interests considering both photographic orientation and angle of view.

In the related literature, some methods to estimate the photograph orientation have been presented, even for photographs

that appear to have no photograph orientation [20], [13]. Our method uses the photographs with photograph orientation to extract area of interests. Therefore, we expect to use the estimated photograph orientation to increase the accuracy of our method.

2.3 Analyzing extracted hotspots

Some researchers study the approach to analyze hotspots obtained from large data annotated with photographic location.

The method to extract hotspots is used to find or detect geographical characterization. Spyrou et al. proposed a method to understand underlying semantics of extracted hotspots using user-generated tags [26]. Omori et al. evaluate geo-referenced photographs annotated with user-generated tags related to coastlines show actual coastline [18]. Hu et al. proposed a method to understand urban areas from extracted hotspots using user-generated tags, and choice preferable photos based on image similarity of photographs in the hotspot [12].

Also, there are some methods to extract the relationship between a hotspot to another hotspot such as the relationship of photographic subjects and shooting spots. Shirai et al. proposed a method to extract a hotspot using DBSCAN and to calculate the relation of hotspots [25]. To discover a wide area of interest, this approach infers the relation among hotspots based on the photograph location and orientation. Hirota et al. proposed a method to detect and visualize various relationship of hotspots using photographic orientation and social tagging [10]. Those researchers extract some relationships from extracted hotspots.

The area of interests extracted by our proposed method represents photographic subjects at which a lot of people took photographs. In other words, comparing to previous methods to extract hotspots, our methods can extract hotspots reflecting interests of users. Therefore, our approach might contribute researchers to characterize the region and extract relationships using hotspots.

3 Proposed method

We propose a method to extract the area of interests using photograph location, orientation, and angle of view obtained from photo-sharing sites. Our approach includes the following steps

1. For a particular area, we obtain vast numbers of photographs from photo-sharing site.
2. Generation of a pseudo triangle of a photograph, which represents photograph area which a photographer was taken, based on photographic location, orientation, and angle of view.
3. From photographs which we use to analyze the particular area, we create grid space and cells which the size specified by a user.
4. we count the number of the pseudo triangles in each cell, and define the number is more than the threshold as the area of interest.

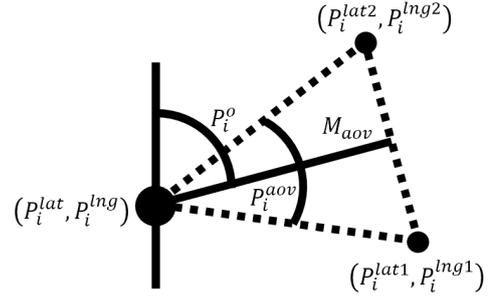


Figure 2: The symbols of pseudo triangle calculated by photographic orientation, location, and angle of view

We describe details related to the respective steps below.

3.1 Calculation of pseudo triangle of a photograph

We describe calculation of a photograph triangle which represents photograph area using photographic orientation, location and angle of view. Here, Figure 2 shows the symbols for following equations to calculate the pseudo triangle of a photograph. The pseudo triangle consists photographic location, orientation, and angle of view.

The angle of view is the angular degree of a given scene that is included in the photograph taken with a camera. The angle of view p_i^{aov} of photograph p_i is calculated from focal length p_i^f and the image sensor size p_i^l as follow.

$$p_i^{aov} = 2 * \tan^{-1} \left(\frac{p_i^l}{2 * p_i^f} \right) \quad (1)$$

To calculate the pseudo triangle of a photograph, we detect the three points of the triangle. The triangle of photograph p_i consists of vertex, angle of the vertex, photographic orientation p_i^o , and the other two points. The vertex of the triangle is the photographic location of latitude (GPSLatitude in Exif) and longitude (GPSLongitude in Exif). The angle of the vertex is angle of view p_i^{aov} of the photograph p_i . The other two points are calculated based on photographic location and orientation p_i^o , the angle of view, and the distance of the photographic subject. We calculate the two points (p_i^{lat1}, p_i^{lng1}) and (p_i^{lat2}, p_i^{lng2}) using following equations.

$$p_i^r = \frac{(p_i^o + p_i^{aov} * 0.5) * \pi}{180} \quad (2)$$

$$d_{lat} = M_{aov} * \cos(p_i^r) \quad (3)$$

$$d_{lng} = M_{aov} * \sin(p_i^r) \quad (4)$$

$$p_i^{lat1} = p_i^{lat} + d_{lat} * \frac{180}{\pi * ER} \quad (5)$$

$$p_i^{lng1} = p_i^{lng} + d_{lng} * \frac{180}{\pi * ER * \cos(\frac{p_i^{lat1} * \pi}{180})} \quad (6)$$

Here, ER is radius of earth. Also, M_{aov} shows the parameter how meter distant from latitude p_i^{lat} and longitude

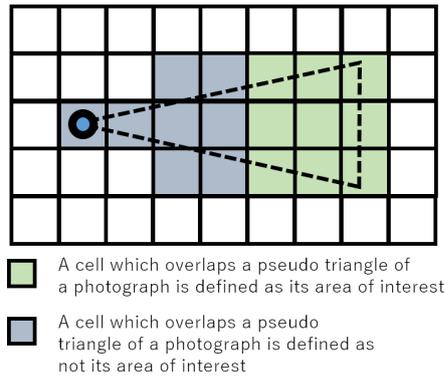


Figure 3: The overlaps of cells and pseudo triangle of a photograph

p_i^{lng} of photographic location of a photo p_i . (p_i^{lat2}, p_i^{lng2}) are calculated by the same procedure that the direction changes $p_i^o + p_i^{aov} * 0.5$. to $p_i^o - p_i^{aov} * 0.5$.

3.2 Calculation of grid

To identify places of area of interest which many people have taken in the photograph, we specifically examined the number of overlaps extracted pseudo triangle of photographs. We create a two-dimensional grid which consists square cells. Here, the area which we want to analyze is much wider than the pseudo triangle of one photograph. Therefore, first, we map photographs which have a photograph location to the grid. Also, using the assigned coordinate of the grid, we count the number of the number of extracted pseudo triangle of photographs.

The photograph p_i are mapped to the coordinate as follows.

$$y_i = M_{height} - \frac{(p_i^{lat} - Lat_{min}) * M_{height}}{Lat_{max} - Lat_{min}} \quad (7)$$

$$x_i = M_{width} - \frac{(p_i^{lng} - Lng_{min}) * M_{width}}{Lng_{max} - Lng_{min}} \quad (8)$$

Here, Lat_{max} , Lat_{min} , Lng_{max} , and Lng_{min} respectively denote the maximum and minimum values of Lat and Lng . Additionally, M_{height} and M_{width} are the height and width of the grid (This is decided using a parameter m for adjustment that how many cells we want to make in this procedure.). Consequently, each cells in the obtained grid include a photograph taken in the range.

3.3 Extraction area of interest

We discover area of interest from the obtained cells using the overlaps of pseudo triangle extracted in previous steps. Figure 3 shows the overlapped cells of pseudo triangle of a photograph. In the figure, the blue circle shows the photographic location. Dashed lines of triangle consist of the pseudo triangle of a photograph. Here, the cells filled with green or gray are overlapped to the pseudo triangle. In this paper, we use the cells filled with green as the users interests shown by a photograph. The reason why our approach does not use some cells close to the photographic location is we think those cells might present the shooting spot, not the interests. Although some cell which might be extracted as

a shooting spot includes many photographs, and the number of overlaps in such cells may be higher than other cells like not a shooting spot. As a result of the case that we do not eliminate the cells filled with gray, almost extracted cells are not the area of interests which our approach wants to extra, but shooting spot. Therefore, in this paper, our approach only uses cell filled with green to extract the area of interest.

We classify whether a cell show interests of a photograph by three cases. First, a cell includes the triangle. Second, the triangle includes a cell. Finally, a line segment of the triangle intersects the line segments of a triangle. Additionally, we remove the cells close to the photographic location, from classified cells.

We apply this procedure to the cells of all photographs and count the number of cells detected as the interests. If the number of photographs in a cell is greater than threshold $minP$, then the cell is classified as an area of interest. Finally, we visualize the extracted area of interest on Google Maps [2].

4 Experiments

This section presents a description of experiments conducted using our proposed method. We present and discuss several examples of detection of visualization of area of interest.

4.1 Datasets

We describe the dataset for experiments for visualizing the area of interests. Photographs for experiments are obtained from photographic search results of Flickr. Those photographs have Exif metaadata of latitude (GPSLatitudeRef, GPSLatitude), longitude (GPSLongitudeRef, GPSLongitude), orientation (GPSImgDirection, GPSImgDirectionRef), and focal length. We obtained 5,842,337 photographs taken during 1 January 2005 - 20 May 2016 and taken in London.

Also, we use compass clustering [15] as comparison method. This clustering method uses the intersections of pseudo orientations of a photograph to extract people's interests. DBSCAN clusters extracted intersections, and the calculated clusters are defined as the area of interests in the method. To simplify comparison between our method and the compass clustering, we change grid-based clustering from the procedure of clustering the intersections. Therefore, we map the intersections into cells of a grid and extract the cells which include the number of photographs than a threshold of $minP$ as the area of interests.

4.2 Visualization of area of interests

Figure 4 presents results of hotspots extracted from photographs in the area of Buckingham Palace, using our proposed method, compass clustering, and shooting spots. We used 2,023 photographs taken in the area (latitude: -0.145 - -0.138 and longitude: 51.506 - 51.499). The clustering parameters are set as $minP = 40$, $M_{aov} = 100$ meters, and $m = 5$ meters. In those figures, the polygon shows the place extracted as a hotspot. In Figure 4b and Figure 4c, the more dark color of a polygon from white to black, the cell includes more numbers of the photographic area of photographs. Also, in Figure 4a, the more dark color of a polygon from white to

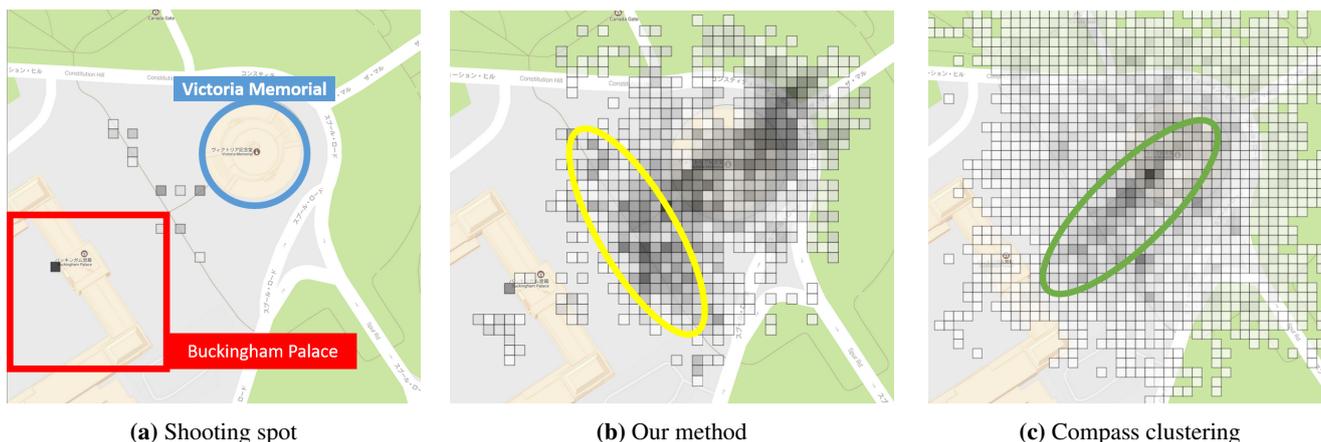


Figure 4: The results obtained from photographs which were taken around Buckingham palace and Victoria Memorial

black, the more photographs were taken in the area. In addition, in those figures, blue circle shows the area of Victoria Memorial, and red rectangle shows the area of Buckingham Palace.

Figure 4a shows the one shooting spot in Buckingham palace, and eleven shooting spots around Victoria Memorial. The shooting spots include the photographs taken of those subjects. However, although the Victoria Memorial is one of famous monument and photographic subject, its place is not extracted as shooting spot. The reason is when people may take a photograph of the subject, they take apart from the subject. On the other hand, in Figure 4b and Figure 4c, the extracted area of interests shows the more wide area than shooting spots of Figure 4a, because those methods use the line of photographic orientations or pseudo triangles.

In result of compass clustering of Figure 4c, there are the dark polygons linearly from Victoria Memorial to Buckingham Palace. Because compass clustering extracts the area of interest using the pseudo orientation of photographs, the method tends to extract the area of interests linearly. On the other hand, in the result of our proposed method in Figure 4b, there are the black polygons widely in the yellow rectangle. Here, the gate of Buckingham Palace exists over a wide range in the area of the rectangle. Therefore, many people take photographs of the gate and Buckingham Palace, and our proposed method discover the place as the area of interests in result shown in Figure 4b. Our approach uses the angle of view to discover the area of interests to consider the region of the interests. As a result, our proposed method can extract the broad subject as the area of interests.

5 Conclusions

We proposed a method to extract and visualize the area of interests using photograph metadata of photographic orientation, location, and angle of view obtained from photo-sharing sites. Our approach identifies a cell in grid mapped from photographs into the area of interest or not, using the overlaps of pseudo triangles extracted by the photograph metadata. We presented some examples of results obtained from Flickr using our proposed method. Comparing to the other method to extract the area of interests based on photographic orientation,

our proposed method can extract the area of interests at more widely, and it is possible to visualize the contents of the area of interests more importantly.

Acknowledgement

This work was supported by JSPS KAKENHI Grant Number 16K00157, 16K16158, and Tokyo Metropolitan University Grant-in-Aid for Research on Priority Areas Research on social big data.

REFERENCES

- [1] Flickr. <http://www.flickr.com/>.
- [2] Google maps. <https://www.google.co.jp/maps>.
- [3] Panoramio. <http://www.panoramio.com/>.
- [4] Rakesh Agrawal, Johannes Gehrke, Dimitrios Gunopoulos, and Prabhakar Raghavan. Automatic subspace clustering of high dimensional data for data mining applications. In *Proceedings of the 1998 ACM SIGMOD International Conference on Management of Data*, SIGMOD '98, pages 94–105. ACM, 1998.
- [5] Mihael Ankerst, Markus M. Breunig, Hans-Peter Kriegel, and Jörg Sander. Optics: Ordering points to identify the clustering structure. In *Proceedings of the 1999 ACM SIGMOD International Conference on Management of Data*, SIGMOD '99, pages 49–60. ACM, 1999.
- [6] Chung-I Chang, Nancy P Lin, and Nien-Yi Jan. An axis-shifted grid-clustering algorithm. *Tamkang Journal of Science and Engineering*, 12(2):183–192, 2009.
- [7] D. Comaniciu and P. Meer. Mean shift: a robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5):603–619, May 2002.
- [8] David J. Crandall, Lars Backstrom, Daniel Huttenlocher, and Jon Kleinberg. Mapping the world's photos. In *Proceedings of the 18th International Conference on World Wide Web*, WWW '09, pages 761–770. ACM, 2009.
- [9] Martin Ester, Hans-Peter Kriegel, Jörg Sander, and Xiaowei Xu. A density-based algorithm for discover-

- ing clusters in large spatial databases with noise. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, KDD'06, pages 226–231, 1996.
- [10] Masaharu Hirota, Motohiro Shirai, Hiroshi Ishikawa, and Shohei Yokoyama. Detecting relations of hotspots using geo-tagged photographs in social media sites. In *Proceedings of Workshop on Managing and Mining Enriched Geo-Spatial Data*, GeoRich'14, pages 7:1–7:6. ACM, 2014.
- [11] Livia Hollenstein and Ross Purves. Exploring place through user-generated content: Using flickr tags to describe city cores. *Journal of Spatial Information Science*, 2010(1):21–48, 2010.
- [12] Yingjie Hu, Song Gao, Krzysztof Janowicz, Bailang Yu, Wenwen Li, and Sathya Prasad. Extracting and understanding urban areas of interest using geotagged photos. *Computers, Environment and Urban Systems*, 54:240–254, 2015.
- [13] Yen-Ta Huang, Kuan-Ting Chen, Liang-Chi Hsieh, Winston Hsu, and Ya-Fan Su. Detecting the directions of viewing landmarks for recommendation by large-scale user-contributed photos. In *Proceedings of the 20th ACM International Conference on Multimedia*, MM '12, pages 997–1000. ACM, 2012.
- [14] Slava Kisilevich, Florian Mansmann, and Daniel Keim. P-dbscan: A density based clustering algorithm for exploration and analysis of attractive areas using collections of geo-tagged photos. In *Proceedings of the 1st International Conference and Exhibition on Computing for Geospatial Research & Application*, COM.Geo '10, pages 38:1–38:4. ACM, 2010.
- [15] Yuri Almeida Lacerda, Robson Gonçalves Fechine Feitosa, Guilherme Álvaro Rodrigues Maia Esmeraldo, Cláudio de Souza Baptista, and Leandro Balby Marinho. Compass clustering: A new clustering method for detection of points of interest using personal collections of georeferenced and oriented photographs. In *Proceedings of the 18th Brazilian Symposium on Multimedia and the Web*, WebMedia '12, pages 281–288. ACM, 2012.
- [16] Xin Lu, Changhu Wang, Jiang-Ming Yang, Yanwei Pang, and Lei Zhang. Photo2trip: Generating travel routes from geo-tagged photos for trip planning. In *Proceedings of the International Conference on Multimedia*, ACM MM'10, pages 143–152. ACM, 2010.
- [17] Amandeep Kaur Mann and Navneet Kaur. Survey paper on clustering techniques. *IJSETR*, 2(4):803–6, 2013.
- [18] Masaki Omori, Masaharu Hirota, Hiroshi Ishikawa, and Shohei Yokoyama. Can geo-tags on flickr draw coastlines? In *Proceedings of the 22nd ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, SIGSPATIAL '14, pages 425–428. ACM, 2014.
- [19] Monali Parikh and Tanvi Varma. Survey on different grid based clustering algorithms. *Int J Adv Res Comput Sci Manag Stud*, 2(2), 2014.
- [20] Minwoo Park, Jiebo Luo, Robert T. Collins, and Yanxi Liu. Beyond gps: Determining the camera viewing direction of a geotagged image. In *Proceedings of the International Conference on Multimedia*, MM '10, pages 631–634. ACM, 2010.
- [21] Adrian Popescu and Gregory Grefenstette. Mining social media to create personalized recommendations for tourist visits. In *Proceedings of the 2Nd International Conference on Computing for Geospatial Research & Applications*, COM.Geo '11, pages 37:1–37:6. ACM, 2011.
- [22] Jörg Sander, Martin Ester, Hans-Peter Kriegel, and Xiaowei Xu. Density-based clustering in spatial databases: The algorithm gdbscan and its applications. *Data Mining and Knowledge Discovery*, 2(2):169–194, 1998.
- [23] Christian Sengstock and Michael Gertz. Latent geographic feature extraction from social media. In *Proceedings of the 20th International Conference on Advances in Geographic Information Systems*, SIGSPATIAL '12, pages 149–158. ACM, 2012.
- [24] Jieming Shi, Nikos Mamoulis, Dingming Wu, and David W. Cheung. Density-based place clustering in geo-social networks. In *Proceedings of the 2014 ACM SIGMOD International Conference on Management of Data*, SIGMOD '14, pages 99–110. ACM, 2014.
- [25] Motohiro Shirai, Masaharu Hirota, Hiroshi Ishikawa, and Shohei Yokoyama. A method of area of interest and shooting spot detection using geo-tagged photographs. In *Proceedings of The First ACM SIGSPATIAL International Workshop on Computational Models of Place*, COMP '13, pages 34:34–34:41. ACM, 2013.
- [26] Evaggelos Spyrou, Michalis Korakakis, Vasileios Charalampidis, Apostolos Psallas, and Phivos Mylonas. A geo-clustering approach for the detection of areas-of-interest and their underlying semantics. *Algorithms*, 10(1):35, 2017.
- [27] Bart Thomee. Localization of points of interest from georeferenced and oriented photographs. In *Proceedings of the 2Nd ACM International Workshop on Geotagging and Its Applications in Multimedia*, GeoMM '13, pages 19–24. ACM, 2013.
- [28] Bart Thomee, Ioannis Arapakis, and David A. Shamma. Finding social points of interest from georeferenced and oriented online photographs. *ACM Trans. Multimedia Comput. Commun. Appl.*, 12(2):36:1–36:23, January 2016.
- [29] Wei Wang, Jiong Yang, and Richard R. Muntz. Sting: A statistical information grid approach to spatial data mining. In *Proceedings of the 23rd International Conference on Very Large Data Bases*, VLDB '97, pages 186–195. Morgan Kaufmann Publishers Inc., 1997.

An approach for e-garment based skill analysis with accelerometer based skill segment extraction

Yu Enokibori**, Shinya Namikawa* and Kenji Mase**

* Graduate School of Information Science, Nagoya University, Japan

**Graduate School of Informatics, Nagoya University, Japan

enokibori@i.nagoya-u.ac.jp, namikawa@cmc.ss.is.nagoya-u.ac.jp, mase@nagoya-u.jp

Abstract - E-Textile based sensing garment is being considered to be used for sports skill analysis because they do not interfere skill motion so much than the other hard-shell type sensors. They are also suitable to collect whole data of game and training, and enables mass-data based skill analysis. However, there is a problem about how to extract segments of target motions. Template matching is not so suitable for new/unknown motion extraction, and also for outlier data extraction, such as data of novice players. Therefore, we propose a new simple approach that extracts segments around significant momentary acceleration feature, such as strong peak, roughly, and then brushes them up. Since most skill motions have acceleration changes, if an accelerometer is attached for the appropriate position, it can generate suitable feature to help find the target motion segment. This approach showed 86.92 % accuracy with e-garment based baseball-throwing-action classification of novice and experienced baseball players.

Keywords: e-textile, e-garment sensor, skill analysis, motion segment extraction

1 INTRODUCTION

Skill analysis based on wearable sensor is being studied actively. These studies are suitable to improve speed of learning, performance of professional athletes, and so on.

Use of e-textile based sensing garment (called e-garment sensor) is a new stream for such studies. E-garment sensor is suitable for this use because it does not interfere body movement. It is an advantage over hard-shell type wearable sensors such as wrist band type accelerometer sensors. Since the characteristics of e-garment sensor, the sensor enables users continuous data collection during games and training situation, and mass-data based skill analysis. For example, if data can be collected during training continuously, effects of sudden coaching can be confirmed through post-process comparisons between pre- and post-coaching data based on life-log style mass-data.

However, such mass-data based skill analysis have an issue: how to extract segments that includes target skill motion. A solution is template-matching based segmentation. However, this approach need a lot of training data. Therefore, it cannot segment not-well-known motions, such as motions of new skills. Moreover, it also cannot segment data including outliers well enough, such as novices' data. Another solution of the issue is automatic segmentation and clustering. However, this approach sometime meets low accuracy

problems and also unnecessary motion segment flood problems.

In order to solve that, we propose a new simple approach that extract segments around momentary acceleration feature, such as strong peak, and then brushes them up. The necessities of this approach are several extra accelerometers that are put on some parts of subjects' body or equipment where do not interfere target skill motions.

We confirmed the potential of the proposed approach with a basic baseball motion analysis and classification. The data of the motion were segmented with two accelerometer, and then were analyzed with an e-garment sensor to measure elbow angle. It showed 83.08 % accuracy for a classification of novice and experienced baseball players with baseball-throwing-action without the brush-up. It improved to 86.92 % accuracy with the brush-up.

The remains of this paper are managed as the followings. In section 2, we describe about related works. In section 3, the details of proposed method is described. In section 4, we discuss about the extraction and brush-up performance of the proposed method. In section 5, the potential of proposed method is evaluated through baseball throwing classification of novice and experienced players with the combination of proposed method and e-garment elbow-angle sensor. Finally, in section 6, we conclude this paper.

2 RELATED WORKS

The followings are summary of related works that separated into sensor based skill analysis and motion segmentation.

2.1 Sensor based skill analysis

E-garment sensors have been studied well, such as LifeShiart [1], Wealthy [2], Myheart [3], MagIC [4] and hitoe [5]. Our projects [6, 7, 8] are also developing several e-textile based sensors including stretch and pressure sensors, and some braid-cord based stretch sensors. Moreover, we are developing several e-garment sensors: for respiration volume measurement, for body pressure measurement and pressure ulcer prevention, and for elbow angle measurement. The advantages of e-textile and e-garment type sensor are flexibility and softness. Thus, they do not interfere skill motion.

Motion and skill analyses based on hard-shell type sensor are studied actively. Okamoto [9] analyzed baseball batting motion with accelerometers and a high-speed camera. Maekawa et al. [10] analyzed manufacturing skill of

assembly lines using accelerometers. A disadvantage of the approaches based on hard-shell type sensor is that sensors sometime interfere skill motions.

2.2 Motion segmentation

Vision based skill-motion segmentations are studied by Archana et al. [11], Chen et al. [12], and so on. Vision based approaches do not interfere skill motions because any wearable sensors are not necessary. However, this approaches have disadvantages of occlusion issues. These also have subject re-identification issues when multiple players captured at once.

Rule and template based motion segmentations are studied by Gageler et al. [13], Blank et al. [14], Murray et al. [15] and so on. Rule and template based motion segmentations do not have disadvantages related to occlusion and subject re-identification. However, these approaches need a lot of training data. Therefore, they cannot segment not-well-known motions, such as motions of new skill. Moreover, they also cannot segment data including outliers well enough, such as novices' data.

Clustering based segmentations are studied by Li et al. [16], Wang et al. [17], Matsubara et al. [18], and so on. These approaches do not need training data. Therefore, they can segment not-well-known motions and data including outliers. However, they sometime meet low accuracy problems and also unnecessary motion segment flood problems.

3 A SIMPLE SKILL SEGMENT EXTRACTION BASED ON ACCELEROMETER

Our approach consists of two phases: the rough segment extraction phase using momentary significant acceleration feature, such as sudden strong peak, and the brush-up phase using dynamic time warping (DTW) distance. The details of them are described in this section.

3.1 The rough segment extraction using momentary significant acceleration feature

Most activities of sports involve acceleration changes on somewhere such as arms, fingertips, and the body itself. For example, tennis swing; baseball throwing, batting and catching; and soccer kick have such characteristic. Moreover, there are several appropriate positions to detect acceleration peaks by acceleration sensors and not interfering skill motions. The necessary of our proposed method is an accelerator attached to one of such positions where subjects select. The flexibility of accelerometer attaching position is an advantage of our proposed method.

An example acceleration data of baseball throwing that is a typical and popular skill to be analyzed is shown in Figure 1. The acceleration data were collected with an accelerometer attached on the wrist of arm used for throwing. The green segment is the true motion segment labeled manually from a ground truth video. We can find a sudden strong peak at about 1.5 seconds. It is included in the true motion segment. This peak is significantly higher than others such as the peaks of right-below part of Figure 1, whose label is running.

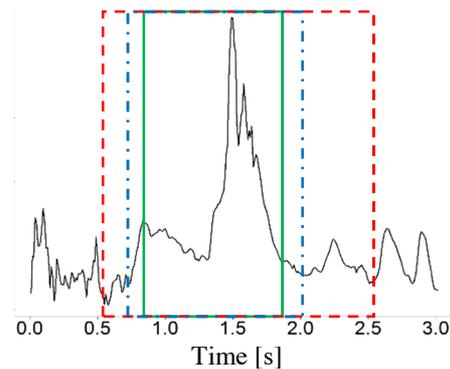


Figure 1: An example of proposed approach (green: true motion segment, red: rough extraction result, blue: brushed-up result)

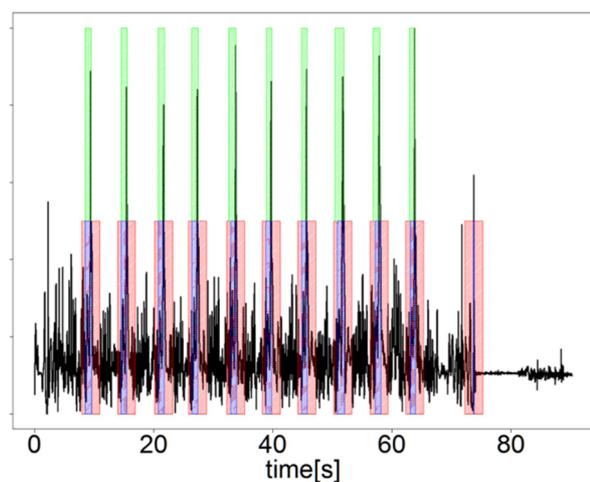


Figure 2: An example of DTW based brush-up (green: true motion segment, red: rough extraction result, blue: brushed-up result)

Therefore, finding of such peaks is easy. Our method finds such peaks and extracts segments around such peaks as rough target motion segments. In the example of Figure 1, we extracted pre and post one second around the peak as a motion segment mentioned with the red dashed box.

3.2 DTW distance based brush-up with simple template

Data segments extracted through the method described in subsection 3.1 include unnecessary data that are not related to target motion. These unnecessary data will lead to several problems, such as low performances of classification, incorrect skill analysis and so on. In order to solve that, our approach have a brush-up phase to remove such unnecessary data segment based on DTW distance.

In contrast with typical DTW based distance calculation, our approach do not need rich template data because the target data of DTW distance calculation are roughly extracted already. Thus, unexpected similar data sequence is not included in most cases. In this study, we built DTW templates from only one motion segment of each motions.

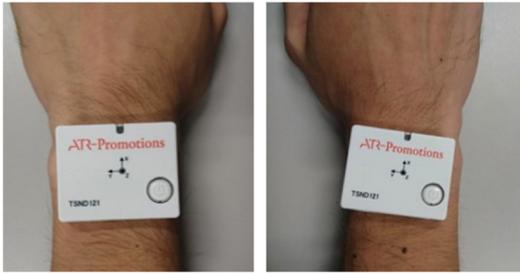


Figure 3: Overview of equipped accelerometer

The brush-up process is the following. First, DTW distances for a template are calculated with sliding and stretching window. For example, if the current target segment contains 100 samples, the window is slid from 1 to 100 and the size of window is changed between 1 and 100. Thus, totally 5050 DTW distances are calculated for all combinations. After that, a segment that has the smallest DTW distance is selected as the most appropriate segment. On the other hand, if the all DTW distances are smaller than a threshold, the segment is removed from extracted segment list. That is perhaps false detection.

An example result of the brush-up is also shown in Figure 1. The blue chain-line segment is the segment brushed up from the red dashed segment. Unnecessary data were removed. The Figure 2 is another example of the extraction and brush-up. The segments brushed up, blue segments, are similar to the ground truth segments, green segments. Moreover, a misdetection nearby 70 second was removed through the brush-up phase.

4 PERFORMANCE DISCUSSIONS ABOUT THE PROPOSED MOTION SEGMENT EXTRACTION AND BRUSH-UP APPROACH

The potential of the proposed motion segmentation and brush-up approach are discussed in this section with an experiment. Baseball throwing and catching motions were used in this experiment because they are typical and popular skill motion to be analyzed.

4.1 Overview of subjects and process

The subjects were six males who were 22 to 24 years old. They equipped accelerometers on their wrists as shown in Figure 3 and collected acceleration data during baseball throwing and catching. The sensors were TSND121 of ATR-Promotions. It worked with 100 Hz sampling rate. Acceleration data of arms used for throwing are used for motion segmentation and brush-up of throwing motion. The opposite arm data are used for segmentation and brush-up of catching motion. Videos also recorded by a camera to build ground truth segments by manually labeling. The overview of experiment process is illustrated in Figure 4. There are two catching points as shown in (a). The details of the process is the following.

1. Moving from a catching point to another one, as (b).
2. Catching a ball thrown from an operator, as (c).

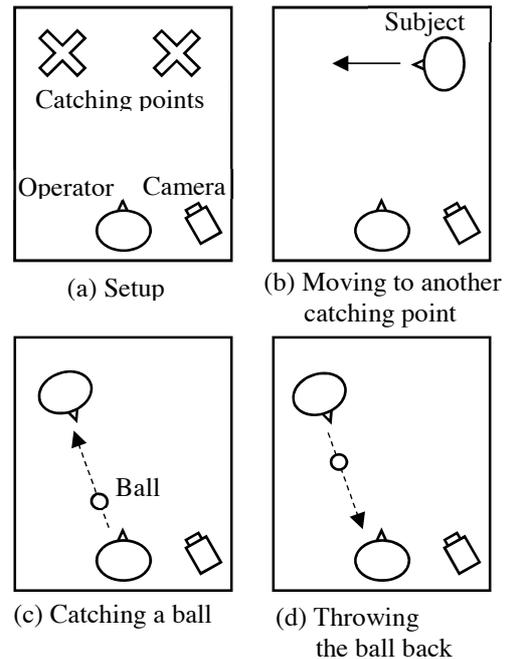


Figure 4: Overview of 1st experiment process

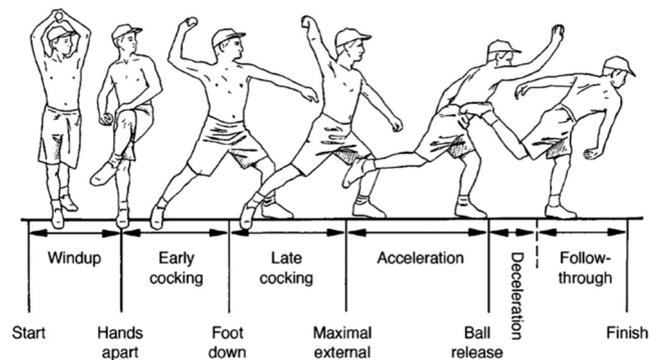


Figure 5: Throwing action phases cited from [26]



Figure 6: Examples of start and finish timing of throwing truth segment

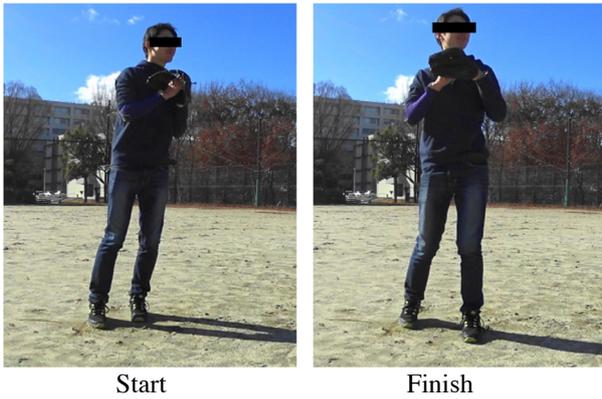


Figure 7: Examples of start and finish timing of catching ground segment

3. Throwing the cached ball for the operator, as (d).
4. Repeating the 1 to 3 actions 10 times.

Therefore, the total number of sampled motion segments is 60.

4.2 Definition of ground truth segment

The ground truth segment definitions of throwing and catching motions used in this study were the following.

4.2.1. Throwing action

Start and finish timings of true segment were defined depending on an example of throwing motion analysis presented in [19] (Figure 5.) As shown in Figure 5, baseball throwing motion is able to be separated into six phases: windup, early cocking, late cocking, acceleration, deceleration, and follow-through. In this study, the start timing of throwing was defined as the initial movement of windup phase: the starting movement of the foot that is opposite side of the arm used for throwing. The finish timing of throwing was defined as the final movement of follow-through: the stopping movement of the arm used for throwing. Examples of the start and finish timings are shown in Figure 6.

4.2.2. Catching action

Example of start and finish timings of catching are shown in Figure 7. The catching action start timing was defined as the starting timing of movement of arm that glove is equipped on. The catching action finish timing was defined as the starting timing of next action, such as holding the cached ball for next throwing.

4.3 Preprocessing for sensor output and parameter selection for proposed approach

The obtained accelerometer data were smoothed with five-point simple moving average, and then normalized for 0 to 1 value-range for each subjects with minimum and maximum values of each subject. 1500 milliseconds was used for the forward/backward extraction time duration around

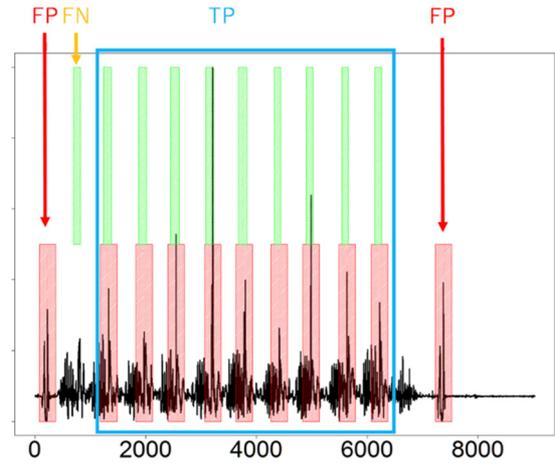


Figure 8: TP/FP/FN examples for extraction score

acceleration peaks. It was defined subjectively depending on the preliminary experiment.

Templates for DTW distance calculation were selected for each subject from the practice phases that were done before the experiment. The DTW distance threshold to remove as misdetection was defined by the following formula.

$$thr = \frac{1}{Num(subjects)} \sum_{s=subjects} thr^s$$

$$thr^s = d_{dtw,mean}^s + w(d_{dtw,med}^s - d_{dtw,min}^s)$$

The d_{dtw}^s is DTW distance of subject s . The $d_{dtw,mean}^s$, $d_{dtw,med}^s$ and $d_{dtw,min}^s$ are mean, median and minimum values of each subject's DTW distances. Finally, the threshold is calculated with weighted sum of each subjects' threshold. The weight w is subjectively defined as 3.2 for throwing and 4.3 for catching depending on preliminary experiment.

4.4 Evaluation methods

The evaluation for our proposed approach was done from two viewpoints: extraction score of motion segments, and overlap score between ground truth and extracted segments.

For simplicity, recall and recall were used with special true positive (TP), false positive (FP) and false negative (FN) definitions described in sections 4.4.1 and 4.4.2. The following typical definitions of precision and recall are used in this study.

$$recall = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

$$precision = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

4.4.1. TP/FP/FN definitions for extraction score

Several TP/FP/FN examples of extraction score are illustrated in Figure 8. The green boxes are ground truth motion segments labeled manually with videos. The red

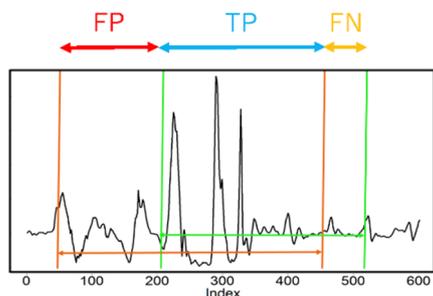


Figure 9: TP/FP/FN examples for overwrap score

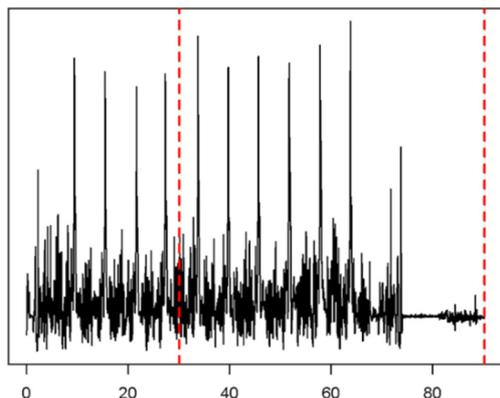


Figure 10: A segmentation result with AutoPlait (red: segmentation points)

boxes are extracted motion segments through our proposed method described in section 3.1. If an extracted motion segment is overwrapped with a ground truth motion segment, it is counted as TP. If there is only a ground truth motion segment, it is counted as FP. If there is only an extracted motion segment, it is counted as FN.

4.4.2. TP/FP/FN definition for overwrap score

TP/FP/FN examples of overwrap score is illustrated in Figure 9. The green segment is a ground truth motion segment labeled manually with videos. The red segment is an extracted motion segment through our proposed method described in section 3.1.

Overwrapped segment of red and green segments is defined as TP. Parts of extracted motion segment where are not overwrapped with any ground truth motion segment are defined as FP. Parts of ground truth motion segment where are not overwrapped with any extracted motion segment are defined as FN.

The unit is sample. With the example of figure 9, since the red and green segments are started at 200 and finished at 450, the TP value is 250.

4.5 Extraction with a related work

Before going into detail discussions, we would like to confirm an extraction result with a state-of-the-art clustering

Table 1: Recall and precision of extraction

	Recall	Precision
Throwing	96.43 ± 8.75	86.80 ± 7.36
Catching	83.26 ± 14.34	63.57 ± 12.98

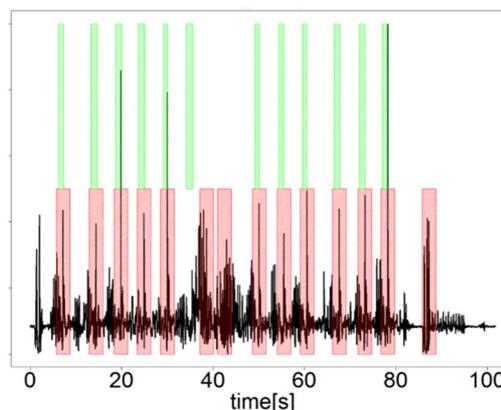


Figure 11: Examples of catching motion extraction

based segmentation method, AutoPlait. The result is shown in Figure 10. This result was calculated with the program code provided in the authors' webpage¹. The red dashed lines are segmentation points provided from AutoPlait. There are only two extracted segments. The same trend, i.e. only two or three extracted segments, were shown with all cases. It is not effectively enough.

It means that automatic segmentation for skill without hints or target limitation is still difficult even if we used state-of-the-art level algorithms. Therefore, we proposed simple and low-cost hints and segmentation method. The following discussions are focused on only the evaluation of our proposed method.

4.6 Extraction score

The extraction scores for baseball throwing and catching actions are summarized in Table 1 with mean and standard deviation values of six subjects. With throwing action, the values of recall and precision were 96.43 ± 8.75 and 86.80 ± 7.36 . The extraction scores for baseball throwing and catching actions are summarized in Table 1 with mean and standard deviation values of six subjects. With throwing action, the values of recall and precision were 83.26 ± 14.34 and 63.57 ± 12.98 . The recall value is worse a little in contrast with the throwing's one. The precision value is significantly worse in contrast with the throwing's one.

An example of extracted catching motion segments are illustrated in Figure 11. As with Figure 8, the green and red segments are ground truth and extracted motion segments. In Figure 11, there is a FN before 40 seconds. Two FPs appears near 40 seconds. Another FP also appears near 90 seconds. The correct label of motions around 40 and 90 seconds were "running." Strong arm swings related to running perhaps yielded strong acceleration changes, and became one of cause

¹ AutoPlait Source Code (visited 2017/07/23): <http://www.cs.kumamoto-u.ac.jp/~yasuko/SRC/autoplait.zip>

of such misdetections. However, the same strong acceleration changes depending on running were also recoded with throwing. Why did not they yield misdetections on throwing extraction? A considerable cause is that operators through the ball too softly in this experiment because all subjects were novices. Therefore, acceleration changes depending on catching became weak, and did not have significant difference between running's one.

4.7 Overwrap score

The overwrap score for baseball throwing and catching actions are summarized in Table 2 with mean and standard deviation values of six subjects. The recall and precision values without proposed brush-up were 96.57 ± 8.41 and 27.80 ± 4.36 for throwing, and 88.11 ± 12.39 and 25.90 ± 9.26 for catching. The not good precisions were expected results because the extraction done roughly as described in section 3.1. These bad precisions were improved with the proposed brush-up: from 27.80 ± 4.36 to 67.06 ± 23.39 with throwing, and from 25.90 ± 9.26 to 63.33 ± 13.32 with catching. On the other hand, recall values were worsened with the proposed brush-up: from 96.57 ± 8.41 to 79.90 ± 15.20 with throwing, and from 88.11 ± 12.39 to 67.61 ± 23.29 with catching.

The worsening of recall does not mean exact failure of proposed method because the aim of this study is improvement of skill analysis with main sensor data. The segmented acceleration data is not used for skill analysis directly. It is have to be discussed whether accuracies of skill analysis with main sensor data will be improved or not. Therefore, we discuss about it in section 5.

5 BASEBALL THROWING MOTION CLASSIFICATION BETWEEN NOVICE AND EXPERIENCED PLAYER WITH E-GARMENT ELBOW-ANGLE SENSOR

As described in last part of section 4, the recall values of overwrap score worsened with proposed brush-up method. However, the worsening of recall does not mean exact failure of proposed method because the aim of this study is improvement of skill analysis with main sensor data. Therefore, with this second experiment, we confirmed whether accuracies of skill analysis with main sensor data will be improved or not.

The second experiment was throwing motion classification of baseball novice and experienced players. Elbow-angles of subjects were measured by an e-garment elbow-angle sensor shown in Figure 12. The collected elbow-angle data were segmented by the proposed method and used for the classification. Details are described as the followings.

5.1 Overview of subjects and process

The subjects were five experienced and seven novice players. All of them were male and 22 to 24 years old. Almost settings and processes were the same as the experiment of section 4. Subjects equipped accelerometers on their wrists as shown in Figure 3 and collected acceleration data during

Table 2: Recall and precision of brush-up

	Recall	Precision
Throwing (w/o)	96.57 ± 8.41	27.80 ± 4.36
Throwing (w/)	79.90 ± 15.20	67.06 ± 23.39
Catching (w/o)	88.11 ± 12.39	25.90 ± 9.26
Catching (w/)	67.61 ± 23.29	86.80 ± 13.32

Meanings of w/ and w/o are with and without brush-up.



Figure 12: E-garment elbow-angle sensor

baseball throwing and catching. The sensors were TSND121 that worked with 100 Hz sampling rate. Videos were also recorded to build ground truth segment with manually labeling. The total number of sampled motion segments was 120.

The difference from the previous experiment was an additional sensor shown in Figure 12. This sensor is an e-garment elbow-angle sensor. The gray lines placed around elbow part are braid-cord stretch sensors that can measure their length change. The sampling rate of this e-garment sensor was 30 Hz.

The baseball throwing motion classification was implemented with support vector machine (SVM) with radial basis function (RBF) kernel. There were no particular reasons to select SVM classifier except the fact that SVM is one of most typical and powerful classifier. The “ksvm” function of “kernlab” package included in R language were used for implementation. The features inputted for SVM are maximum, minimum, mean and variance values of elbow angle. Time duration of segment and appearance timing of acceleration peak were also inputted. The appearance timing was normalized from 0 to 1 values that mean start and end of motion segment.

5.2 Evaluation methods

Performances of extraction and brush-up were evaluated with the same manner described in section 4.4. If details are necessary, please refer section 4.4.

Performances of the baseball throwing motion classification were evaluated with one-subject-out style training and test. One subject data were used as test data and other data were used as training data. Accuracies were calculated from total results of all combinations. This experiment is a binary classification. Thus, the formula of accuracy is the following.

$$\text{accuracy} = \frac{\text{True Positive}}{\text{Total Sample Number}}$$

Table 3: Recall and precision of extraction with the experiment of section 5

	Recall	Precision
Experienced	100.00 ± 0.00	96.67 ± 7.45
Novice	100.00 ± 0.00	90.48 ± 13.11

Table 4: Recall and precision of brush-up with the experiment of section 5

	Recall	Precision
Experienced (w/o)	100.00 ± 0.00	40.00 ± 6.23
Experienced (w)	91.98 ± 5.91	88.49 ± 10.17
Novice (w/o)	100.00 ± 0.00	31.23 ± 5.74
Novice (w)	77.80 ± 12.46	82.66 ± 14.36

Meanings of w/ and w/o are with and without brush-up.

5.3 Extraction and overlap score

The extraction score of this experiment is summarized in Table 3. Both values of recall and precision with this experiment showed higher performance than the results of the previous experiment. A cause of this high performance is a change of throwing policy. In section 4, the operators and the subjects threw balls so softly because all of the subjects were novice. On the other hand, in this experiment, they threw balls with normal speed. Then, acceleration peaks became more significantly. It perhaps one of cause of the high performance.

The overlap score of this experiment is summarized in Table 4. The both recall values without proposed brush-up showed so high performance. On the other hand, the precision values without proposed brush-up were not so good. These were only 40.00 ± 6.23 and 31.23 ± 5.74 for experienced subjects and novice subjects. It is the same trend with the result of previous experiment described in section 4. Such worse precisions were improved with the proposed brush-up: from 40.00 ± 6.23 to 88.49 ± 10.17 with experienced subjects and from 31.23 ± 5.74 to 82.66 ± 14.36 with novice subjects. In contrast with the improvement, recall values were worsened from 100.00 ± 0.00 to 91.98 ± 5.91 with experienced subjects and from 100.00 ± 0.00 to 77.80 ± 12.46 with novice subjects. This trend is also the same with the result of the previous experiment.

As described above, there is no significant trend difference with the previous experiment although there are several improvement depending on the change of throwing policy.

5.4 Accuracies of baseball throwing motion classification between novice and experienced players with e-garment elbow-angle sensor

The accuracies of the baseball-throwing-motion classification with the e-garment elbow-angle sensor between novice players and experienced players are summarized in Tables 5 and 6. The accuracies with ground truth segments are summarized in Table 7.

The total sampling numbers of Tables 5 and 6 is bigger than the ground truth segments number, 120. It is because the

Table 5: Accuracies of baseball throwing motion classification between novice and experienced players without brush-up

Label	Estimated		Accuracy
	Exp.	Nov.	
Experienced	35	17	67.31
Novice	5	73	93.59
Total			83.08

Table 6: Accuracies of baseball throwing motion classification between novice and experienced players with brush-up

Label	Estimated		Accuracy
	Exp.	Nov.	
Experienced	37	15	71.16
Novice	2	76	97.44
Total			86.92

Table 7: Accuracies of baseball throwing motion classification between novice and experienced players with ground truth segments

Label	Estimated		Accuracy
	Exp.	Nov.	
Experienced	37	13	74.00
Novice	1	69	98.57
Total			88.33

total number of Tables 5 and 6 includes misdetections of proposed extraction method.

As shown in Tables 5 and 6, the accuracies were improved with the proposed brush-up in all cases: in experienced players, in novice players, and also total. There is 3.84 points (4.6 %) improvement between the total accuracies. Moreover, the improved total accuracy have 98.4 % performance of ground truth based accuracy. Therefore, the proposed method showed a potential to be suitable to skill motion data collection and segmentation.

6 CONCLUSION

In this paper, we proposed a new simple approach to extract motion data segmentation that extracts segments around significant momentary acceleration feature, such as sudden strong peak, and then brush them up. The results of extraction showed high recall and precision scores: 96.43 ± 8.75 and 86.80 ± 7.36 for baseball throwing motion, and 83.26 ± 14.34 and 63.57 ± 12.98 for baseball catching motion. On the other hand, the result of brush-up was not significantly good.

However, the not-so-good result of brush-up does not mean exact failure of proposed method because the aim of this study is improvement of skill analysis with main sensor data. Therefore, we confirmed whether accuracies of skill analysis with main sensor data will be improved or not by proposed approach with the second experiment: baseball throwing motion classification between novice and experienced players with an e-garment elbow-angle sensor.

The result of the second experiment showed high total accuracy, 86.92 %, for the classification with the proposed brush-up. It is 3.84 points (4.6 %) improvement from the accuracy, 83.08 %, without the brush-up. Moreover, the improved total accuracy showed 98.4 % performance of ground truth based accuracy, 88.33 %. It means that the proposed method showed a potential to be suitable to skill motion data collection and segmentation.

ACKNOWLEDGEMENT

This research was partly supported by JSPS KAKENHI Grant Number 15H02736 and 26280074, and “The Knowledge Hub of AICHI, The Priority Research Project”.

REFERENCES

- [1] F.H. Wilhelm, W.T. Roth, and M.A. Sackner. The LifeShirt: An Advanced System for Ambulatory Measurement of Respiratory and Cardiac Function. *Behavior Modification*, Vol. 27, No. 5, pp. 671-691, 2003.
- [2] R. Paradiso, G. Loriga, and N. Taccini. A Wearable Health Care System Based on Knitted Integrated Sensors. *IEEE Transactions on Information Technology in Biomedicine*, Vol. 9, No. 3, pp. 337-344, 2005.
- [3] M. Pacelli, G. Loriga, N. Taccini, and R. Paradiso. Sensing Fabrics for Monitoring Physiological and Biomechanical Variables: E-textile solutions. In *Proceedings of the 3rd IEEE-EMBS*, pp. 1-4, 2006.
- [4] M.D. Rienzo, P. Meriggi, F. Rizzo, P. Castiglioni, C. Lombardi, M. Ferratini and G. Parati. Textile Technology for the Vital Signs Monitoring in Telemedicine and Extreme Environments. *IEEE Transactions on Information Technology in Biomedicine*, Vol. 14, No. 3, pp. 711-717, 2010.
- [5] K. Takagahara, K. Ono, N. Oda and T. Teshigawara. 'hitoe' - A Wearable Sensor Developed through Cross-industrial Collaboration. *NTT Technical Review*, Vol. 12, No. 9, pp. 1-4, 2014.
- [6] Y. Enokibori, A. Suzuki, H. Mizuno, Y. Shimakami and K. Mase. SpiroVest: An e-Textile-Based Wearable Spirometer with Posture Change Adaptability. In *Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication*, pp. 203-206, 2013.
- [7] Y. Enokibori, A. Suzuki, H. Mizuno, Y. Shimakami and K. Mase. E-textile pressure sensor based on conductive fiber and its structure. In *Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication*, pp. 207-210, 2013.
- [8] Y. Enokibori, T. Hayashi and K. Mase. A Study of Intermittent Adjustment to Resist Displacement of Smart Garment using Posture-stable Daily Actions. In *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers*, pp. 249-252, 2015.
- [9] S. Okamoto. Evaluation of Skills in Batting Technique in Baseball Using Acceleration Sensors. *Journal of the Japanese Society for Experimental Mechanics*, Vol. 12, No. Special_Issue, pp. s255-s260, 2012.
- [10] T. Maekawa, D. Nakai, K. Ohara and Y. Namioka. Toward Practical Factory Activity Recognition: Unsupervised Understanding of Repetitive Assembly Work in a Factory. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pp. 1088-1099, 2016.
- [11] M. Archana and M. Kalaisevi Geetha. Object Detection and Tracking based on Trajectory in Broadcast Tennis Video. *Procedia Computer Science*, Vol. 58, pp. 225-232, 2015.
- [12] W. Chen and Y. Zhang. Tracking Ball and Players with Applications to Highlight Ranking of Broadcasting Table Tennis Video. In *Computational Engineering in Systems Applications, IMACS Multiconference on*, Vol. 2, pp. 1896-1903, 2006.
- [13] W.H. Gageler, S. Wearing and D.A. James. Automatic jump detection method for athlete monitoring and performance in volleyball. *International Journal of Performance Analysis in Sport*, Vol. 15, No. 1, pp. 284-296, 2015.
- [14] P. Blank, J. Hoßbach, D. Schuldhuis and B.M. Eskofier. Sensor-based Stroke Detection and Stroke Type Classification in Table Tennis. In *Proceedings of the 2015 ACM International Symposium on Wearable Computers*, pp. 93-100. ACM, 2015.
- [15] N.B. Murray, G.M. Black, R.J. Whiteley, P. Gahan, M.H. Cole, A. Utting and T.J. Gabbett. Automatic Detection of Pitching and Throwing Events in Baseball with Inertial Measurement Sensors. *International Journal of Sports Physiology and Performance*, pp. 1-18, 2016.
- [16] L. Li, J. McCann, N.S. Pollard, and C. Faloutsos. DynaMMo: Mining and Summarization of Coevolving Sequences with Missing Values. In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 507-516. ACM, 2009.
- [17] P. Wang, H. Wang and W. Wang. Finding Semantics in Time Series. In *Proceedings of the 2011 ACM SIGMOD International Conference on Management of Data*, pp. 385-396. ACM, 2011.
- [18] K.J. Jones, D.C. Osbahr, M.A. Schruppf, J.S. Dines and D.W. Altchek. Ulnar Collateral Ligament Reconstruction in Throwing Athletes: A Review of Current Concepts. *J Bone Joint Surg Am*, Vol. 94, No. 8, p.e49, 2012.

Keynote Speech 2:
Prof. Ivica Botički
(Faculty of Electrical
Engineering and Computing,
University of Zagreb)

IWIN 2017 Conference, Zagreb

Learning with Gamification, Collaboration and Augmented Reality Technology



Prof. Ivica Botički, PhD

Faculty of Electrical Engineering and Computing, University of Zagreb, Croatia

Soolla^M

Prof. Ivica Botički, PhD

- *Faculty of Electrical Engineering and Computing, University of Zagreb, Croatia*
- *Teaching: Algorithms and Data Structures, Programming Paradigms and Languages, Object Oriented Programming and other PhD courses*
- *Research: Technology-enhanced learning, mobile learning, information systems*

Soolla^M

Three Challenges in Technology-Enhanced Learning Research

- *Challenge 1: Design innovative technology for learning*
- *Challenge 2: Implement technology in educational settings*
- *Challenge 3: Evaluate the impact of technology*

Sōolla[™]

3

About the STOLLAm Project

- *Croatian National Science Foundation
2014-2017*
- *STOLLAm = Opening up education
through Seamless and COLLABorative
Mobile learning on tablet computers:*
 - *Gamified learning experiences*
 - *Augmented reality learning experiences*
 - *Computer supported collaborative learning*



Sōolla[™]

4

SOllAm Technology System Architecture and the Author System



Soolla^M

5

DEMO 1: The Author System

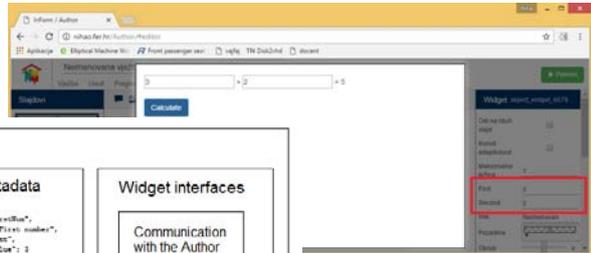
- <http://nihao.fer.hr/Author>



Soolla^M

6

Widgets as Reusable Digital Educational Content Units



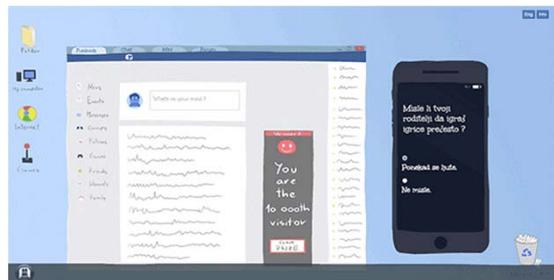
Widget (adding two numbers)

<p>Widget code (web-based)</p> <pre><html ng-app="myApp"> <head> <scriptes.... </head> <body> <div class="container" ng-controller="MainCtrl"> <input type="number" name="first" ng-model="first"> + <input type="number" name="second" ng-model="second"> = {{ result }} <button class="btn btn-primary" ng-click="submit()"> Calculate/buttons </div> </body> </html></pre>	<p>Widget metadata</p> <pre>"interface": [{"name": "firstNum", "label": "First number", "type": "text", "defaultLabel": "1"}, {"name": "secondNum", "label": "Second number", "type": "text", "defaultLabel": "2"}], "returnValue": [{"name": "myData", "label": "log data", "type": "object", "defaultLabel": "{}"}] }</pre>	<p>Widget interfaces</p> <ul style="list-style-type: none">Communication with the Author systemGroup communication via Socket.IO
---	--	--



First thing First: Safety on the Internet

- <http://www.sofolla.com/Application/PeCoApp/2>



DEMO 2: Gamification Technology for Learning

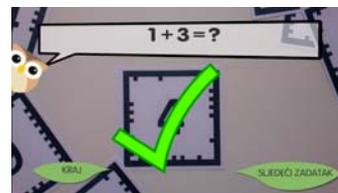
- <http://www.scollam.com/Application/VirusApp/13>



Soolla^M

DEMO 3: Augmented Reality Technology for Learning

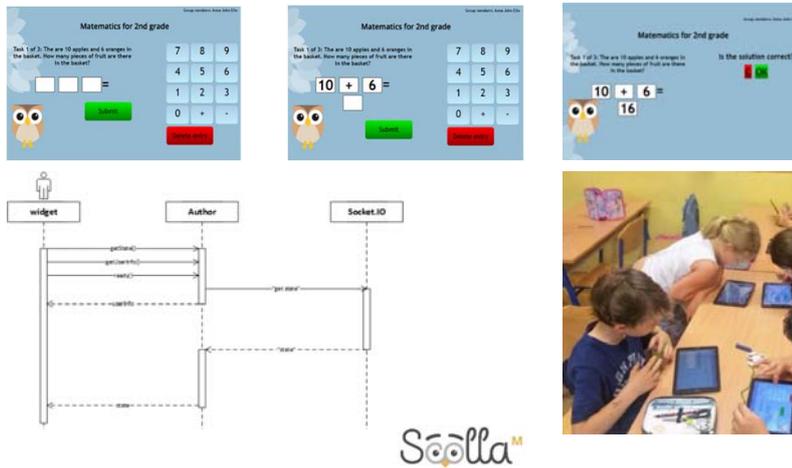
- <http://www.scollam.com/Application/ARApp/15>



Soolla^M

DEMO 4: Collaboration Technology for Learning

- <http://www.scollam.com/Application/CollaborationApp/14>



11

Research Findings: Benefits of Gamification for Learning

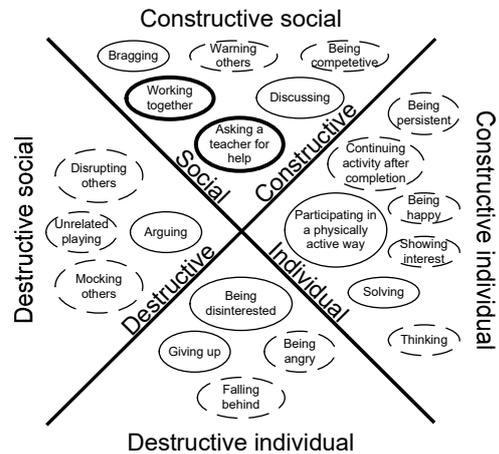
- *Early primary school students were engaged when they were challenged at a suitable difficulty level*
- *Engagement drops when the level of challenge gets too high and students get frustrated.*
- *The level of student engagement can be greatly enhanced with the careful combination and integration of different game elements, such as narratives and adaptive mechanism based on individual performance*

Soolla^M

12

Research Findings: Benefits of Augmented Reality Learning

- *The results indicate that students participating in more static ARLEs exhibit more constructive social actions and positive emotional engagement*
- *Students participating in more mobile ARLEs exhibit more constructive individual actions and cognitive engagement.*



Research Findings: Benefits of Collaborative Learning

- *Work on data analysis still in progress*
- *On a macro level, the preliminary analysis shows that students, even first-graders, are able to successfully engage in synchronous collaborative math learning.*
- *Preliminary results indicate the need for a more adaptive group and role assignment mechanisms to better support learners in different tasks and stages of their learning.*

www.scollam.com



Thanks!

Ivica.boticki@fer.hr

Scolla[™]

15

Session 7:
Network and Security
(Chair: Yoh Shiraishi)

A Scheduling Algorithm for Slotted-CSMA Based Wireless Mesh Networks with High-speed Links

Yasuhiro Mori^{†*} and Takuya Yoshihiro^{‡**}

[†]Graduate School of Systems Engineering, Wakayama University, Japan

[‡]Faculty of Systems Engineering, Wakayama University, Japan

* s181056@sys.wakayama-u.ac.jp

** tac@sys.wakayama-u.ac.jp

Abstract - In this paper, we propose to extend CATBS, which is a slotted-CSMA-based architecture for wireless mesh networks (WMNs), to support high-speed IEEE802.11 links. In CATBS, a single frequency channel is divided into several large slots in which CSMA runs to avoid collisions, and the schedule that assigns slots to links eliminates collisions due to so called hidden terminal problem. However, since the scheduling method of CATBS is based on single-disk interference model, the schedule would not work well when high-speed link is used in WMNs. This paper proposes to use the double-disk interference model in CATBS in the scheduling algorithm to enable CATBS to use high-speed links.

Keywords: Wireless Mesh Networks

1 INTRODUCTION

Although wireless Mesh Networks (WMNs) has been deeply studied so far [1], the performance in the current state of the art is far from practical due to heavy interference coming from hidden terminal effects. In order to reduce the hidden terminal effects over MAC of IEEE802.11 [2], RTS/CTS handshake is traditionally used [3]. However, several studies showed that RTS/CTS does not work well in WMNs due to physical property of wireless communications [4][5][6]. Even recently, IEEE802.11-based WMNs have not been improved to achieve practically sufficient performance [7].

As an architecture to achieve practical performance, we have proposed the scheme called CATBS (CSMA-Aware Time-Boundable Scheduling). CATBS is based on slotted CSMA, i.e., CSMA works within time-divided slots, and avoid collisions due to hidden-terminals by applyin a schedule that assigns a slot for each link. CATBS theoretically eliminates collisions and robust to the time drift on synchronizing time among nodes. However, one of the problems is that CATBS originally adopts single-disk interference model so that it suffers from severe interference with high-speed communication links because the schedule does not match the physical property of communications. In this paper, we propose a new scheduling method for CATBS that takes high-speed links into account. Specifically, we adopt double-disk interference model in the schedule computation to reflects the physical property of high-speed links on the schedule.

In the following, we first present CATBS in Sec.2. In Sec. 3, we present the scheduling method based on double-disk interference model. In Sec. 4 we evaluate the proposed method, and finally we conclude the work in Sec.5.

2 CATBS: A Slotted-CSMA-Based Architecture of WMNs

2.1 Overview

CATBS is a communication method for WMN that avoids hidden terminal problem, which is achieved with a combination of slotted CSMA and a scheduling method. In addition, the slotted CSMA used in CATBS is a little modified from the original CSMA. First, a single frequency channel is time-divided to create several virtually independent channels. Then, CSMA runs within each virtual channel. In the scheduling method, a virtual channel is allocated to each link such that hidden terminal problem does not occur. Since collision between adjacent nodes can be avoided by the carrier-sensing function of CSMA, our scheduling method only take the hidden-terminal effects into account. We define the interference model that considers radio interference due to hidden terminal problem and formulate it as an optimization problem that minimizes the effect of hidden terminal problem. Since the formulated problem is proven to be NP-hard, we obtain the solutions efficiently by reducing to PMAX-SAT.

2.2 MAC Protocol

The MAC protocol used in CATBS is a slotted CSMA with a little modification. The slotted CSMA is a mechanism in which we divide a frequency channel with a fixed time interval and run CSMA within each slot. Different from the original slotted CSMA, we do not use TDMA in slots at all because TDMA requires strict time synchronization. By avoiding to use TDMA, CATBS works with relatively loose time synchronization system, which significantly relaxes the restriction for the communication system.

In the MAC protocol used in CATBS, first, a single frequency channel is divided in time and multiple virtual channels are generated. Each virtual channel is called a slot. Then, CSMA runs inside the created slot. In order to operate CSMA, it is necessary to take a relatively large time per slot compared with TDMA. Each slot is given a number $1, 2, \dots, k$ for identifying them, and it is switched in turn as $1, 2, \dots, k, 1, 2, \dots$ and so on. An example is shown in the Fig. 1. In Fig. 1, a single frequency channel is time-divided into k slots. Since two links in the relationship of hidden terminal problem transmit frames in different slots, the hidden terminal problem does not occur. In addition, RTS / CTS is used to avoid frame collision at the boundary of slots due to time synchronization

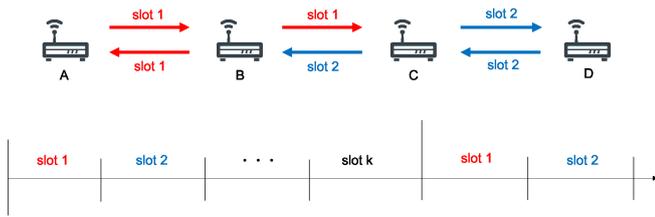


Figure 1: Virtual multi-channelization by time division

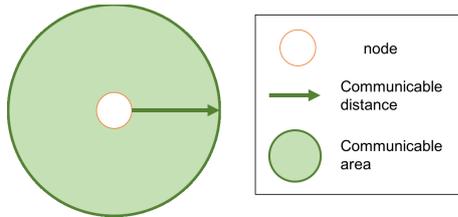


Figure 2: Communicable distance

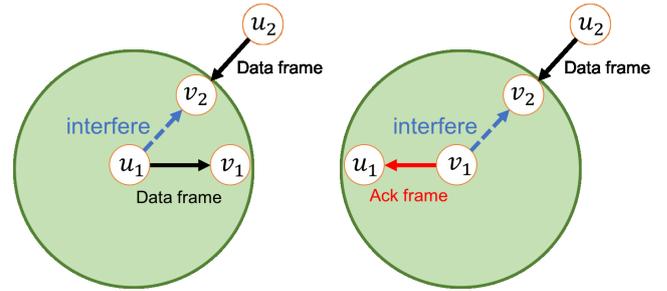
error. Namely, when transmitting a data frame, if the transmission time is judged to exceed the boundary of the current slot, RTS is transmitted. After CTS is returned, nodes that received RTS or CTS wait for NAV period without transmitting data frames even if when the allocated slot comes. When the NAV period ends, the nodes start transmitting data frames.

2.3 Definitions

In order to formulate the scheduling problem, we begin with definitions. The network is represented by the directed graph $G = (V, E, C)$, where V is the node set, E is the link set, and C is the channel set. We define $e = (u, v, c)$ as the link to communicate using channel c from node $u \in V$ to $v \in V$. If there is a pair of links $e_1 = (u_1, v_1, c_1)$ and $e_2 = (u_2, v_2, c_2)$ in the hidden terminal relationship, it is called as an interference link pair. We denote the shortest path length from node u to v on graph G by $D_{(u,v)}^G$. Let S_G be the set of interference link pairs in G , which is called as the collision degree of graph G . The scheduling problem of CATBS aims at minimizing the number of collision link pairs $|S_G|$ by removing links in G and output the graph G' that is free from hidden terminal problem.

2.4 Single-disk Interference Model

CATBS uses a single disk model as an interference model to simplify the situation where radio interference occurs. In the single disk model, when a node communicates with some other nodes, the distance with which the communication succeeds is called the communicable distance r , and the area inside the circle of radius r is called the communicable area. In the single disk interference model, we assume that there is no radio interference outside the communicable area.



(a) Data frame interferes

(b) Ack frame interferes

Figure 3: conditions of interference link pairs

2.5 Defining Collision Link Pairs Based on Single Disk Model

Collisions under single disk model are modeled as two types: collisions invoked by data frames, and those invoked by Ack frames. We show an example of those two types of collisions in Fig. 3. In Fig. 3(a), a data frame from u_1 to v_1 collides with a frame from u_2 to v_2 . In Fig. 3(b), an Ack frame sent from v_1 collides with a frame from u_2 to v_2 . Those two types of collisions are formally defined as follows.

Type 1: collision with data frames occur if all the following conditions are met.

1. $c_1 = c_2$
2. $(u_1, u_2, c_1) \notin E$
3. $(u_1, v_2, c_1) \in E$

Type 2: collision with Ack frames occur if all the following conditions are met.

1. $c_1 = c_2$
2. $(u_1, u_2, c_1) \notin E$
3. $(v_1, v_2, c_1) \in E$

2.6 Formulation of Scheduling Problem

In the scheduling problem formulation of CATBS, we first consider the graph G that consists of every possible links $e = (u, v, c)$ where $u, v \in V$ and $c \in C$, i.e., we include links with every combinations of u, v, c . Then, we choose a subset of links in G and output the schedule $G' = (V, E', C)$ where $E' \subseteq E$. Note that, from the restriction of the default router architecture that each node has only one transmission queue, the number of slots assigned to a node is limited to one. Namely, the incoming and the outgoing links of the same node must belong to the same slot. Also note that the pair of links in the relationship of collision has already defined. Our goal is to minimize the *interference level*, which is defined as the number of collision link pairs in G' .

If the number of slots in the schedule increase, CATBS would face in severe end-to-end delay due to the time to wait active slot at each node. To prevent the delay in scheduling, CATBS allow to use paths that are not the shortest-paths between the source and destination node pairs. Specifically, in

the scheduling, CATBS chooses a set of links for G' such that, for each pair of source and destination (s, d) , the length of the shortest path in G' is equal to or less than that in $G + k$, where k is the predefined stretch factor. More formally, if we let $\delta_{s,d}^G$ be the shortest-path length from s to d in G , $\delta_{s,d}^{G'} \leq \delta_{s,d}^G + k$. Namely, by allowing k -hop longer paths than the shortest path, CATBS reduces the number of slots required to achieve zero-collision. As above, the formal description of the scheduling problem in CATBS is given as follows.

The Scheduling Problem in CATBS

Input: A graph $G = (V, E, C)$, A set of collision link pairs S_G .

Output: A schedule $G' = (V, E', C)$ where $E' \subseteq E$

Subject to: $\delta_{s,d}^{G'} \leq \delta_{s,d}^G + k$, and every node does not have more than 2 assigned slots.

Minimize: Interference level $|S_{G'}|$

2.7 Algorithm to Solve the Problem

In the literature the scheduling problem was proven to be NP hard. Therefore, it takes a huge amount of time to find the optimal solution. In CATBS, in order to find an approximate solution efficiently, we reduce the scheduling problem to PMAX-SAT. PMAX-SAT is a traditional NP-hard optimization problem, and recently, there held a contest of good solvers for large-scale PMAX-SAT problems, for which several excellent solvers have been developed so far. CATBS intends to use one of those excellent solvers of PMAX-SAT.

In PMAX-SAT, we let x_1, x_2, \dots, x_n be logical variables that take values true (1) or false (0). Let \bar{x}_1 be the inverted value of logical variable x_1 . A logical expression such as $(x_1 \vee x_2)$ obtained by connecting several logical variables with OR operators (\vee) is called a clause. We call logical expressions in which we connect clauses with AND operator (\wedge) as a canonical normal form (CNF) formulas, e.g., $(x_1 \vee x_2) \wedge (\bar{x}_1 \vee x_3)$. For each of the logical variables x_1, x_2, \dots, x_n in the given CNF formula we assign a logical value true(1) or false(0). The SAT(SATisfiability Problem) problem is defined as a task to output whether there is a set of true/false assignment that satisfy the given CNF formula. The problem to maximize the number of satisfied clauses is called MAX-SAT (MAXimum SATisfiability problem). As a further extension of MAX-SAT, we define Partial MAX-SAT (PMAX-SAT). For a given CNF formula $f(x_1, x_2, \dots, x_n) = g_h(x_1, x_2, \dots, x_n) \vee g_s(x_1, x_2, \dots, x_n)$ where $g_h(\cdot)$ and $g_s(\cdot)$ are also CNF formula called hard and soft clauses, respectively, PMAX-SAT maximize the number of satisfied clauses in soft clauses g_s under the constraint that all the hard clauses g_h are satisfied. The formal description of PMAX-SAT is as follows.

PMAX-SAT

Input: CNF formula $f = g_h \vee g_s$.

Output: 0/1-Assignment of logical variables.

Maximize: The number of satisfied soft clauses.

Subject to: All hard clauses are satisfied.

We make a reduction from the scheduling problem of CATBS to PMAX-SAT. The constraints of CATBS such as the increase of path length are expressed by the hard clauses, and optimization criterion, i.e., the interference level, is expressed by the soft clauses. Specifically, for the input of the scheduling problem $G = (V, E, C)$, we define logical variables $l_{u,v,c}$ for all links included in E , where $l_{u,v,c}$ takes true if the corresponding link exists in G' and false otherwise. Although we omit the detail due to paper limitation, the hard clauses g_h are created such that they all are satisfied only if all the constraints in the scheduling problem are satisfied. See reference [11] for detail. The soft clauses g_s consists of a set of clauses $(\bar{l}_1 \vee \bar{l}_2)$, each of which corresponds to each collision link pair (l_1, l_2) in G . This clause does not satisfy only if both links are included in G' and invoke collision. As a result, once the PMAX-SAT is solved, the set of binary variables $l_{u,v,c}$ determines the schedule G' , with which the interference level $|S_{G'}|$ is minimized. When this formula is a hard clause and the hard clause is true, the graph G' satisfies the constraints of the optimization problem. Next, in the soft clause, the logical expression $(\bar{l}_1 \vee \bar{l}_2)$ for all the link pairs included in the set of link pairs S'_G in the hidden terminal problem relationship Take it with an AND operator. $(\bar{l}_i \vee \bar{l}_j)$ takes false if both link pairs in hidden terminal problem relationships are not restricted on graph G' . That is, the number of logical expressions that take false matches the collision degree on graph G' . Then, by allocating logical variables that takes as many true as possible in soft clauses, graph G' with lowest collision degree is output.

2.8 The Problem with CATBS-based WMNs

CATBS uses the single disk model to simplify the situation of hidden terminal problem. However, wireless communication gets more vulnerable against noise when the communication speed gets higher. Specifically, with high-speed links, the communicable distance goes shorter while the interference distance stays the same. As a result, when we assume high-speed links, schedules based on the single-disk model are no more possible to treat collisions appropriately. From this reason, we in this paper introduce the double-disk interference model to compute schedules more suitable for high-speed wireless communications.

3 The Proposed Method

3.1 overview

We propose a new scheduling method to reduce radio interference in WMNs with high speed links. In our proposal, we use a double disk model as a more realistic interference model than the single disk model. In the double disk model, two distances, i.e., the communicable distance and the interference distance are defined such that two nodes can communicate with each other if they are located within the communicable distance, but a radio from a node located within the interference distance disturbs it. Our scheduling method is designed for distributed environment. First, we describe the method

to identify the nodes within the range of the interference distance. We next give the algorithm to compute the collision link pairs from that information. After those steps, we can compute the schedule based on the double disk model.

3.2 Concept of proposed method

We intend to realize autonomously decentralized networks in which each node executes schedules and forwarding paths by means of running routing protocols. In the schedule calculation, each node requires the information of the network topology, so that a routing protocol such as OLSR is used to perform the distributed control. In order to realize the autonomous distributed scheduling using a routing protocol, each node needs to grasp the node located in the interference area under the double disk model. However, since the interference distance is larger than the communication distance especially with high-speed links, this information cannot be grasped by control messages of routing protocols. Also from the proactive routing protocol, each node cannot grasp the distance information to identify the nodes in the interference distance because this kind of routing protocols only treats the topology of the network.

To grasp the nodes within the interference area, we use beacons that is periodically transmitted by every node with low communication speed. By observing all beacons received at every node, it grasps a set of nodes from which beacons can be received with high probability. Each node regards that the set of nodes are within the interference area. Since beacons are specified to transmit with minimum possible speed, the interference distance is supposed to be far larger than communicable distance.

In order to perform the distributed control using routing protocols, it is necessary that the calculation time of the schedule should be short and schedules should be computed even on a terminal with low capability. Simultaneously, we must design a joint routing and scheduling protocol in which schedules and routing tables are surely computed and the routing scheme works. In this paper, we simply introduce a feasible design of the joint routing and scheduling protocol that works in the distributed environment, and propose the method that can incorporate with the protocol.

3.3 Routing and Scheduling Protocol Framework

In order to execute the proposed scheduling method autonomously and distributedly, each node must collect the information required for scheduling. To compute schedules, we must collect two sorts of information, i.e., the network topology, and a set of collision link pairs. The network topology can be collected by using a proactive routing protocol such as OLSR. Therefore, we in this paper describe only how to collect collision link pairs.

In our joint routing and scheduling protocol, each node first computes its schedule (as we mentioned before, a schedule G' is a subgraph of the network topology G), and computes the shortest paths on G' to build its routing table. To compute a schedule, we must collect the network topology and the set of

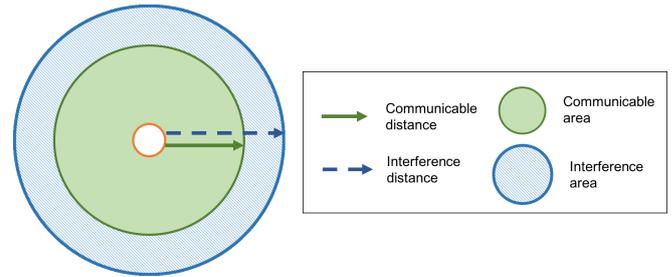


Figure 4: Interference distance

collision link pairs. To compute the latter, each node observes beacons and grasps a set of nodes within the interference area. On the other hand, the nodes within the communicable distance is known from the network topology. By sharing those two sorts of information with the surrounding nodes, each node can compute the collision link pairs. The set of collision link pairs computed at every node is shared over the network, and all nodes in the network obtains the set of collision link pairs as a result. Since every node know the topology and the collision link pairs of the whole network, a joint routing and scheduling protocol as described above is able to design. As the framework we apply the proposed scheduling algorithm, we assume this kind of network protocols.

3.4 Defining Collision Link Pairs Based on Double Disk Model

In the proposed method, we apply double disk model as the interference model to take the radio interference under high-speed links into account. The double disk model is defined as two disks with different radius, which represents the communicable area and the interference area, respectively. In this interference model, the communicable distance (i.e., radius) means that two nodes are communicable with each other if other radio does not exist. Similarly, the interference distance means that the communication from a node s to d fails if d is within the interference distance from a node i on which transmission is ongoing. Generally, in high-speed links, the communicable distance is far smaller than the interference distance.

In order to formulate the scheduling problem, we make definitions of collision link pairs. For the network represented by a directed graph $G = (V, E, C)$, we assume two links as $e_1 = (u_1, v_1, c_1)$ and $e_2 = (u_2, v_2, c_2)$, and define the condition in which e_1 interferes e_2 due to the hidden terminal effect. Here, we denote the set of nodes located in the interference area of u_1 by N_{u_1} .

As in the case of CATBS, collisions are classified into two patterns: one is the case where data frames collide to other frames, and the case Ack frame collides. An example is shown in Fig. 5. In Fig. 5(a), the data frame from u_1 to v_1 collides to another frame from u_2 to v_2 . In contrast, In the case of Fig. 5(b), the Ack frame from v_1 to u_1 collides to another frame from u_2 to v_2 . The formal representation of those two type of collision cases are written as follows.

Type 1: collision with data frames occur if all the following

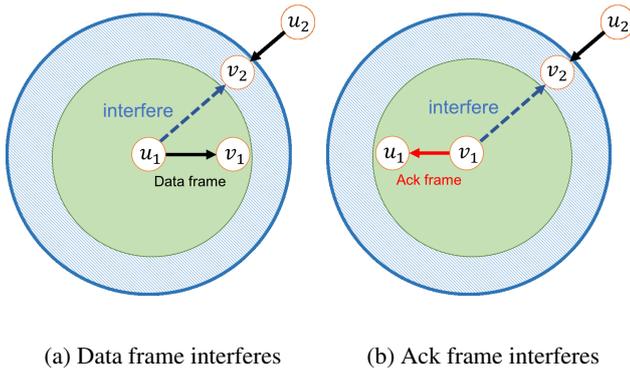


Figure 5: Conditions of interference link pairs

conditions are met.

1. $c_1 = c_2$
2. $u_2 \notin N_{u_1}$
3. $v_2 \in N_{u_1}$

Type 2: collision with Ack frames occur if all the following conditions are met.

1. $c_1 = c_2$
2. $u_2 \notin N_{u_1}$
3. $v_2 \in N_{v_1}$

4 Evaluation

4.1 Evaluation method

We evaluate the proposed scheduling method through simulation using network simulator Scenargie [9]. In the simulation, we clarify whether the proposed scheduling method enables efficient communication with few frame collisions under high speed communication. For the simulation, we used a 5×5 grid topology in which each node can communicate directly with vertically and horizontally adjacent nodes. The communication speed is set at 36 Mbps, 48 Mbps, which is high speed communication in IEEE 802.11g. The distance between each node was set to 370[m] from the preliminary test mentioned above. In our scenario, 12 CBR (Constant Bit Rate) communication flows are generated vertically and horizontally as shown in Fig. 6(a). We show the schedule used in Fig. 6(b). The packet size is 1500[Bytes] and the transmission power is 20[dbm]. We start measuring the performance at 60 seconds after the flows are generated. We measured the packet delivery rate, the total throughput, and the delivery delay to evaluate the traffic performance.

4.2 Simulation Results

Fig. 7(a)-(c) and Fig. 8(a)-(c) show the total throughput, packet delivery ratio, and delivery delay, respectively. Comparing the throughput performance, the proposed method maintains high throughput in both cases of 36 Mbps and 48 Mbps. In the case of 48[Mbps], difference gets larger as the transmission rate increases, and when transmission rate is 4000

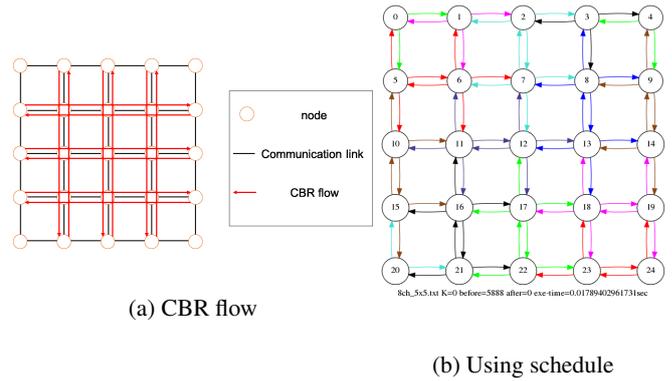


Figure 6: Experiment scenario

[kbps], the difference comes as much as about 1.5 times. The difference between CATBS and the proposed method comes from the suitability of schedules, i.e., it means that the double-disk model is more suitable to fit to the property of high-speed links.

The packet delivery ratio shows the similar trend as the throughput performance. The proposed method keeps more than 90 % in both cases of 36 and 48[Mbps]. Especially, in the case of 48 [Mbps], the delivery rate of CATBS gradually decreases from the early stage, while the 36[Mbps]-case the performance is kept until the transmission rate reaches 3400[Kbps]. This trend shows that the higher-speed communication, i.e., 48[Mbps], is more vulnerable against noise, and radio interference that the schedule cannot cover increases. Nevertheless, the proposed method actually works well even in the 48[Mbps]-case to reduce the interference effect from communication performance.

When we look at delivery delay to the destination, we see that the delay arises in CATBS from the lower transmission rate than the proposed method. Note that a large delivery delay means saturation of networks, i.e., the transmission queue of nodes overflowed due to the effect of interference. We see that the delivery delay also has the same trend as throughput and delivery ratio. This means that the major factor to reduce communication performance is interference: the effect of interference reduces the network capacity, and when the amount of traffic exceeds the capacity, all the performance criteria, i.e., throughput, delivery ratio, and delivery delay, reduces at the same time. Note that this phenomenon occurs more gradually in the 48[Mbps] case while the performance degradation occurs rapidly in the 36[Mbps] case.

Based on above, we conclude that the proposed method can suppresses radio interference under high-speed communication and maintains high communication performance as compared with CATBS. However, even with the proposed method, the effect of interference gets larger in the 48[Mbps] case, and as a result, the total performance of the 36[Mbps] case is better than the 48[Mbps] case. From the communication log of the simulator, in the 48[Mbps] case, frame collisions caused by radio far from the place of packet failure are found. Namely, even small-level noise could degrade the performance of communication in high-speed links. Our beacon-based model parameter determination works well for 36[Mbps] links, but causes degradation in the 48[Mbps] case.

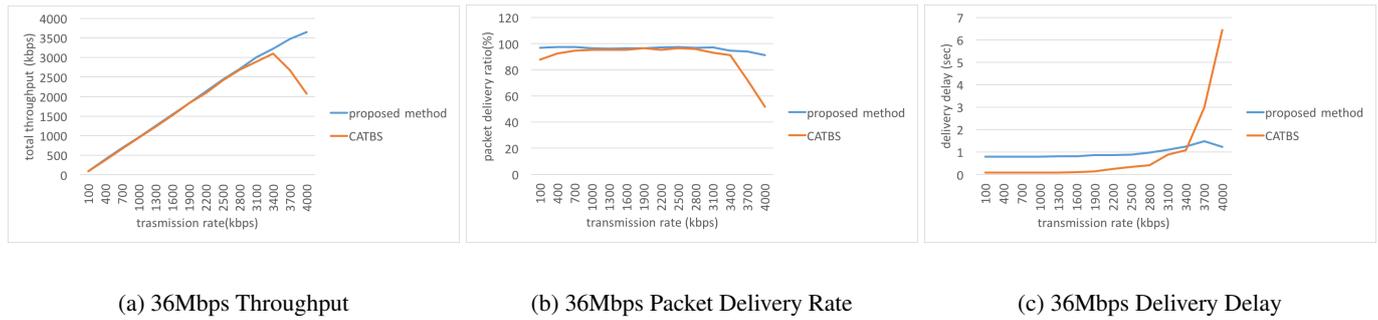


Figure 7: 36Mbps Simulation result

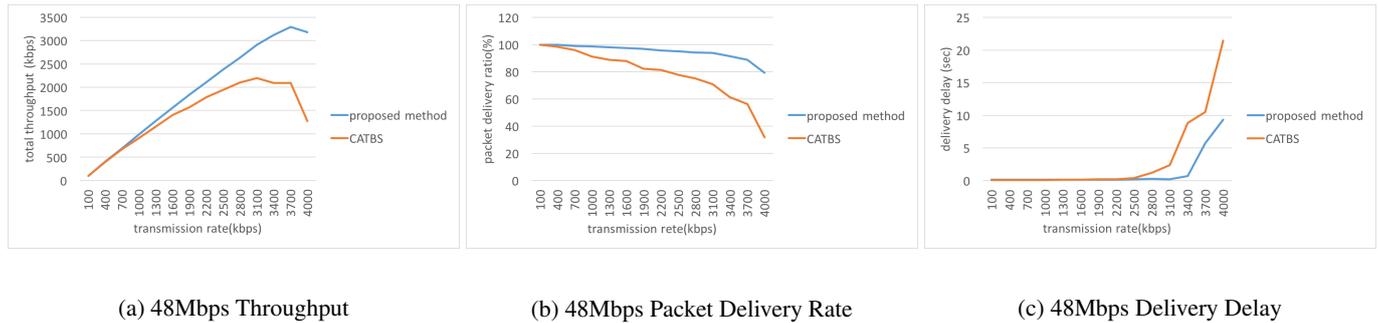


Figure 8: 48Mbps Simulation result

5 Conclusion

In this paper, we proposed a new scheduling method to reduce radio interference under high speed communication. By using the double disk model as an interference model, more precise treatment of radio interference is possible compared to the conventional single disk model. The proposed scheduling algorithm based on the double disk model, a large part of collisions that are involved in the schedule in CATBS are avoided.

As a result of the evaluation, we confirmed that, by using the proposed scheduling method, radio interference under high speed communication is reduced and communication performance is improved. From this fact, modeling using the double disk model is more suitable than using the single disk model under high-speed communication. In addition, the proposed method uses beacon reception status to determine the interference distance. Our evaluation in this paper shows that we can grasp the interference distance with this approach, and indicates the possibility that we could design autonomous distributed joint routing and scheduling scheme in the future.

One of the future task is to grasp the optimal topology for the proposed method by conducting simulation experiments with various topologies. Also, note that the communicable distance and the interference distance may vary according to communication speed. Therefore, it is necessary to confirm how effective the method to determine the interference distance with beacons is. Furthermore, in the simulation, radio interference which cannot avoid by the proposed method was observed. In order to prevent such radio interference, it is also an issue to consider more realistic interference models.

Acknowledgment

A part of this work is supported by KAKENHI(16K12422).

REFERENCES

- [1] Akyildiz, I. and Wang, X.: Wireless Mesh Networks, John Wiley & Sons, Ltd., Publication, 2009.
- [2] IEEE802.11 Wireless local Area Networks, <http://www.ieee802.org/11/> (referred in Feb 2017).
- [3] B. Bharghavan et al., "MACAW: A Media Access Protocol for Wireless LANs," In Proc. ACM SIGCOMM'94, 1994.
- [4] J.L. Sobrinho, R. de Haan, J.M. Brázio, "Why RTS-CTS Is Not Your Ideal Wireless LAN Multiple Access Protocol," In Proc WCNS'05, 2005.
- [5] K. Xu, M. Gerla, and S. Bae, "Effectiveness of RTS/CTS Handshake in IEEE 802.11 Based Ad Hoc Networks," Ad Hoc Networks, Vol.1 Issue.1, pp.107-123, 2003.
- [6] Xu, K. et al.: Effectiveness of RTS/CTS Handshake in IEEE802.11 Based Ad Hoc Networks, Ad Hoc Networks, Vol.1, No.1, pp.107-123, 2003.
- [7] R.K. Sheshadri and D. Koutsonikolas, "Comparison of Routing Metrics in 802.11n Wireless Mesh Networks," The 32nd IEEE International Conference on Computer Communications (INFOCOM'13), 2013.
- [8] Takuya Yoshihiro and Taiki Nishimae, "Practical Fast Scheduling and Routing over Slotted CSMA for Wireless Mesh Networks," In Proc. IWQoS2016, 2016.
- [9] Network Simulator Scenargie, Space Time Engineering, <https://www.spacetime-eng.com/jp/> (referred in Jan 2017).

A Strategy for Mitigating Information Loss in Anonymization by Using Nature of Personal Data

Takayasu Yamaguchi^{†‡} and Hiroshi Yoshiura[‡]

[†]Research Laboratories, NTT DOCOMO, Inc.

[‡]Department of Informatics, University of Electro-Communications
yamaguchitaka@nttdocomo.com, yoshiura@uec.ac.jp

Abstract - Statistics generated from a collection of personal data are used for various purposes such as marketing, recommendation, and urban planning. However, research has shown that there is a risk of the personal data being inferred from the statistics. To avoid such inference and protect the privacy of the individuals represented by the statistics, anonymization is used to modify the statistics. Such modification, however, degrades their accuracy. The greater the number of an individual's attributes in the statistics, the easier it is to infer his/her personal data and thus the larger the modifications needed to maintain his/her privacy. On the other hand, the fewer the number, the less useful the statistics. We propose optimizing this trade-off by making use of the dependency between personal attributes. Our strategy is to directly use some of the attributes in the personal data while indirectly using the remaining ones by estimating their values from the values of the attributes used. The relationships between the used and remaining attributes are used for this estimation. For example, given the attributes of height, weight, and gender, height and gender are directly used and weight is estimated using the relationship between the three attributes, thus reducing the amount of modification by a third while maintaining the usefulness of the statistics. However, if the relationships themselves are sensitive, i.e. reflecting the personal data, additional anonymization is needed, which could spoil the strategy. Thus, the key to making our strategy effective is to identify the relationships between personal attributes that are precise and insensitive. The effectiveness of our strategy was demonstrated by implementing it for differential privacy, a representative anonymization technique, and by evaluating the implemented system using the MovieLens 1M dataset.

Keywords: security, big data, data mining, anonymization, differential privacy

1 INTRODUCTION

Statistics generated from a collection of personal data are used for various purposes such as marketing, recommendation, and urban planning. However, research has shown that there is a risk of the personal data being inferred from the statistics. To prevent such inference and protect the privacy of the individuals represented by the statistics, anonymization is used to modify the statistics. Such modification, however, degrades their accuracy. The greater the number of an individual's attributes in the statistics, the easier it is to infer his/her personal data and thus the greater the number of modifications needed to maintain his/her privacy. On the other

hand, the fewer the number, the less useful the statistics. We propose optimizing this trade-off by making use of the relationships between personal attributes.

Our strategy is to use some of the attributes in the personal data while indirectly using the remaining ones by estimating their values from the values of the attributes used. The relationships between the used and the remaining attributes are used for this estimation. For example, given the attributes of height, weight, and gender, height and gender are used and weight is estimated using the relationship between the three attributes, thus reducing the amount of modification by a third while maintaining the usefulness of the statistics. However, if the relationships themselves are sensitive, i.e. reflecting the personal data, additional anonymization is needed, which could spoil the strategy. Thus, the key to making our strategy effective is to identify the relationships between personal attributes that are precise and insensitive.

Implementation of our strategy depends on the anonymization method used. A promising next-generation framework for anonymization is differential privacy, and a representative method for implementing differential privacy is the Laplace mechanism [1]. We have developed an anonymization method based on the Laplace mechanism and evaluated it using the MovieLens 1M dataset, a standard dataset for evaluating the performance of recommendation systems. We estimated recommendation accuracy by comparing recommendation performance based on the original dataset with that based on the anonymized dataset.

2 RELATED WORK

2.1 Anonymization methods

Anonymization is applied either to records of personal information (called microdata) or to statistics calculated from microdata. The main purpose of anonymizing microdata is to prevent linking records to specific persons and that of anonymizing statistics is to prevent inferring the original microdata from which the statistics were derived. Representative methods for anonymizing microdata are methods based of k-anonymity [2], l-diversity [3], and t-closeness [4]. A set of techniques generically called statistical disclosure control are used for anonymizing statistics [5][6]. Methods based on differential privacy [1] and probabilistic k-anonymity [7] are used for both. Anonymization is implemented by modifying data, and methods for modification include perturbing, swapping, aggregating, and omitting data [8]. Thus, it is inevitable that anonymization degrades the original data and statistics. Two

essential issues of anonymization are therefore security and data quality [9]. Security is related to how well the purpose of anonymization (i.e. preventing linkage and inference) can be achieved. Data quality is related to how well the information that can be extracted from the data or statistics is preserved. Because there is a trade-off between security and data quality, our goal is to improve security while maintaining the quality or to improve quality (i.e. minimize information degradation) while maintaining security.

2.2 Differential Privacy and Laplace Mechanism

Differential privacy is a criterion for achieving anonymization security that is particularly promising due to its rigid mathematical basis. It is expressed as an inequality:

$$Pr[\mathcal{K}(D_1) \in S] \leq \exp(\epsilon) \times Pr[\mathcal{K}(D_2) \in S], \quad (1)$$

where $\mathcal{K}(D_x)$ is a mechanism for calculating and anonymizing statistics from microdata D_x , and ϵ is a security parameter. D_1 and D_2 are microdata that are different in only one record. Thus, equation (1) represents that the ratio between two anonymized statistics is limited in the range of exponential (ϵ) if they were derived from two sets of microdata that differ in only one record (i.e. differ for only one person). The smaller the ϵ , the greater the anonymization.

The Laplace mechanism, a representative mechanism for implementing differential privacy, adds Laplacian noise with a scale parameter (λ) to the statistics. The scale parameter is sometimes referred to as the diversity. The larger the λ , the greater the noise and the more degraded the statistics. The Laplace mechanism achieves differential privacy by satisfying an inequality:

$$\lambda \geq \frac{\Delta f}{\epsilon}, \quad (2)$$

where ϵ is the security parameter mentioned above, and Δf is the sensitivity of the statistical value against the microdata. It is the maximum change in the statistics when one record (i.e. one person) in the microdata is changed. The smaller the Δf , the better the quality of the anonymized statistics. This is reasonable considering the meaning of differential privacy mentioned above.

2.3 Issues in Quality Maintenance

The loss of quality due to using the Laplacian mechanism can be decreased by reducing Δf , which depends on the type of statistic. For example, it is proportional to $1/n$ for average and variance, where n is the number of records in the microdata, and it is independent of n for maximum and minimum. Thus, average and variance suffer less degradation due to using the Laplacian mechanism than maximum and minimum. One approach to maintaining quality is thus to reduce the target statistics to a combination of statistics that have good properties such as average and deviation [10]. Though this approach has received a lot of attention, the conventional differential privacy mechanism is not practical for many applications.

By ‘‘sensitivity’’ we mean the maximum change in a statistic when any possible change in one record is considered in the context of any possible combination of other records [11] [12] [13]. Thus, another approach to reducing Δf is to ignore rare cases [14][15][16] that make Δf large. This approach, however, degrades anonymization security because anonymization does not work if rare cases occur.

The fewer the attributes of a record (i.e. of a person), the lower the sensitivity [17] [18]. For example, assume a record has three attributes: gender, height, and weight. The statistics to be calculated are the frequencies of tall men (men taller than 180cm) and heavy men (men heavier than 80kg). The sensitivity is 2 because a man of 190cm and 90kg could change into a woman of 150cm and 50kg¹. If we consider only gender and height, the sensitivity is reduced to 1 because we can count only the frequency of tall men. Thus, a third approach is to consider a subset of the record attributes. However, straightforward methods in this direction simply use less information to calculate the statistics [19], which is less useful for such applications as marketing of sporting goods. Using primary component analysis (PCA) is one way to reduce the number of attributes while maintaining information quality [20][21][22]. However, PCA may reveal private information in itself, e.g. how the values of the attributes are related.

3 Strategy for Reducing Number of Attributes

We maintain the quality of the anonymized statistics by using the third approach mentioned above; i.e. we use a subset of attributes in the microdata to calculate the statistics. We maintain the quality of information by directly using a subset of the attributes and estimating the values of the other attributes by using the relationships between the directly used attributes and the other attributes. We derive the relationships between the attributes from public information, thereby preventing the revelation of private information. If we cannot derive the relationships from public information only, we derive them by using statistics for the microdata such as the average and deviation, which are less sensitive than the target statistics. That is, we calculate insensitive statistics, which suffer less from anonymization, anonymize them, use the anonymized statistics to establish relationships between attributes, use these relationships to estimate the values of the unused attributes, calculate the target statistics, and finally anonymize the target statistics with less degradation.

We explain our strategy by using the example shown in Figure 1. The microdata are shown at the top left. The target statistics are the frequencies of tall men and heavy men which are represented by a frequency table. A conventional method to obtain an anonymized frequency table is to generate a frequency table and anonymize it. However, the sensitivity in this case is 2 (as mentioned above), and may be too high to maintain the quality of the anonymized statistics. We thus use only values of gender and height in the microdata (at upper right of Figure 1) and generate a frequency table from the reduced microdata. We anonymize this reduced frequency table, which reduces the sensitivity to 1.

¹This example is a little bit complex than that described in the abstract.

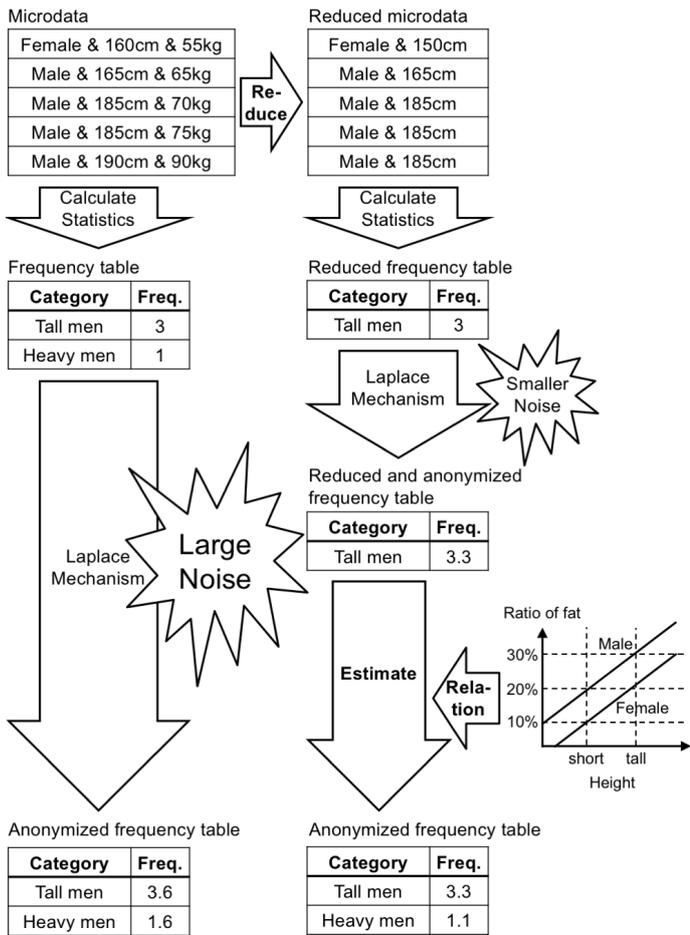


Figure 1: Example implementation of proposed strategy

We then estimate the frequency of heavy men by using the relationships between gender, height, and weight. Ideally, these relationships would be precisely obtained from public information such as statistics published by the Ministry of Health, Labour and Welfare. Otherwise, we can obtain the relationships as a combination of insensitive statistics such as average and deviation.

4 Example Implementation and Evaluation

4.1 Implementation

Our use case is the construction of a recommendation system for movies. The system finds the top ten movie categories that are best suited for the age, gender, and occupation of the user. A post-processing system selects movies from those categories [23] [24].

We used the MovieLens 1M dataset [25] as example microdata for this system. The dataset consists of 1,000,209 records, each of which is a user's rating of a movie. It contains data for 6,040 users and 3,952 movies. Table 1 shows a part of this dataset. Note that there are 18 basic movie categories such as Action, Adventure and Comedy, and there are

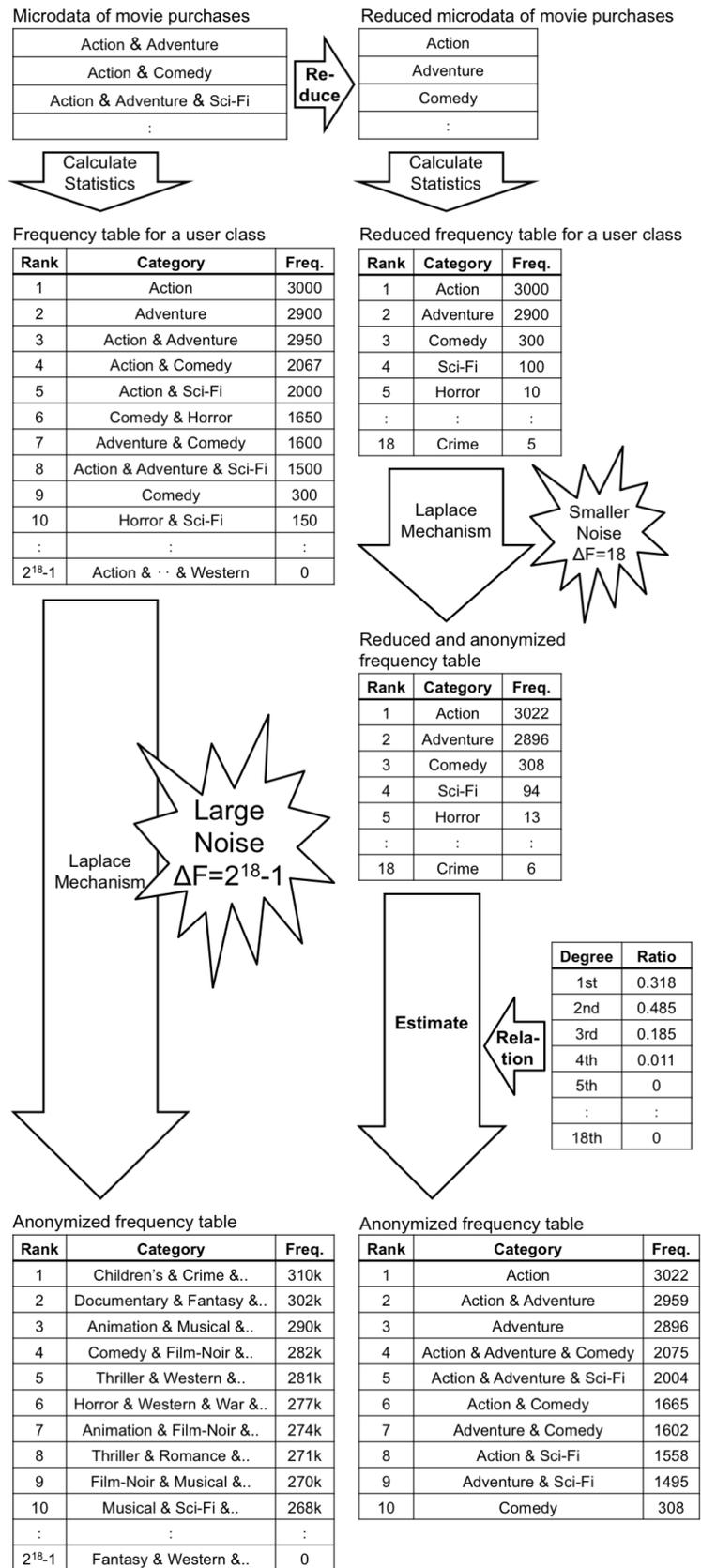


Figure 2: Overview of example implementation

Table 1: Part of MovieLens 1M dataset

User ID	Attribute	Movie (category)	Rating
1	Male, 18–24, programmer	Waterworld, (Action & Adventure)	5
1	Male, 18–24, programmer	Beverly Hills Cop, (Action & Comedy)	4
2	Female, 25–34, writer	Sabrina, (Comedy & Romance)	3
3	Male, 18–24, programmer	Star Trek, (Action & Adventure & Sci-Fi)	5
4	Female, 25–34, artist	Sound of Music, (Musical)	4

Table 2: Results for using gender and age

Required Security Level	Without Anonymization	Conventional Method	Reduced-Attributes Method (Order)			Proposed Method
			1st	2nd	3rd	
None	100%	100%	28%	78%	99%	48%
Low	NG	0%	28%	51%	1%	48%
Moderate	NG	0%	28%	34%	0%	48%

$2^{18} - 19$ combinatorial categories such as “Action & Adventure”, “Comedy & Romance”, and “Action & Adventure & Sci-Fi”². Categories combining two basic categories, e.g. “Action & Adventure”, are called second-order categories, those combining three are called third-order categories, and, in general, those combining i basic categories are called i -th-order categories. Thus, we have $2^{18} - 1$ basic and combinatorial categories. The user ratings range from 1 to 5, with 5 being the highest.

Records with rates larger than four were considered to be for movies purchased by the users. The microdata of purchases thus obtained were used to generate a frequency table for each user class (e.g. “Female & 25-34 & writer”). The frequency table is shown in the upper-left part of Figure 2. It was anonymized, as shown at the bottom-left. The top ten categories in the anonymized table were recommended by the system with respect to the user class. In the anonymization, however, the sensitivity was $2^{18} - 1$ because one person could purchase all $2^{18} - 1$ categories or purchase none of them.

To reduce the sensitivity, we used only 18 basic categories in line with the strategy explained in Section 3 and estimated the frequencies for the other $2^{18} - 19$ categories using simple relationships of averaging frequencies for basic categories, e.g. frequency of “Action & Adventure” is (frequency of Action + frequency of Adventure) / 2 and that of “Action &

Table 3: Results for using occupation

Required Security Level	Without Anonymization	Conventional Method	Reduced-Attributes Method (Order)			Proposed Method
			1st	2nd	3rd	
None	100%	100%	30%	82%	100%	50%
Low	NG	0%	30%	37%	2%	49%
Moderate	NG	0%	30%	18%	1%	46%

²There are $2^{18} - 19$ combinations of basic categories except for the basic categories themselves and a null combination is excluded.

Table 4: Results for using gender, age, and occupation

Required Security Level	Without Anonymization	Conventional Method	Reduced-Attributes Method (Order)			Proposed Method
			1st	2nd	3rd	
None	100%	100%	28%	77%	93%	47%
Low	NG	0%	26%	10%	1%	36%
Moderate	NG	0%	24%	8%	1%	27%

Table 5: Detailed results for using gender and age ($\epsilon = 2.0$)

Gender	Age	Reduced-Attributes Method (Order)			Proposed Method
		1st	2nd	3rd	
Female	-18	30%	0%	0%	50%
	18–24	30%	20%	0%	50%
	25–34	30%	40%	10%	50%
	35–44	30%	40%	0%	50%
	45–49	30%	30%	0%	60%
	50–55	30%	0%	0%	50%
	56–	20%	0%	0%	40%
Male	-18	30%	0%	0%	40%
	18–24	30%	60%	0%	50%
	25–34	30%	70%	0%	50%
	35–44	20%	60%	0%	40%
	45–49	30%	30%	0%	50%
	50–55	30%	40%	0%	50%
	56–	30%	40%	0%	50%
Weighted Average		28%	51%	1%	48%

Table 6: Detailed results for using gender and age ($\epsilon = 1.0$)

Gender	Age	Reduced-Attributes Method (Order)			Proposed Method
		1st	2nd	3rd	
Female	-18	30%	0%	0%	50%
	18–24	30%	10%	0%	50%
	25–34	30%	20%	0%	50%
	35–44	30%	10%	0%	50%
	45–49	30%	10%	0%	60%
	50–55	30%	0%	0%	50%
	56–	20%	0%	0%	30%
Male	-18	30%	0%	0%	20%
	18–24	30%	40%	0%	50%
	25–34	30%	60%	0%	50%
	35–44	20%	30%	0%	40%
	45–49	30%	20%	0%	50%
	50–55	30%	10%	0%	50%
	56–	30%	30%	0%	50%
Weighted Average		28%	34%	0%	48%

Table 7: Detailed results for using occupation ($\epsilon = 2.0$)

Job	Reduced-Attributes Method (Order)			Proposed Method
	1st	2nd	3rd	
academic / educator	30%	40%	0%	50%
artist	30%	30%	0%	50%
clerical / admin	30%	10%	0%	50%
college / grad student	30%	70%	0%	50%
customer service	30%	0%	0%	50%
doctor / health care	30%	10%	0%	50%
executive / managerial	30%	70%	10%	50%
farmer	20%	10%	0%	20%
homemaker	20%	10%	0%	50%
K-12 student	30%	10%	0%	50%
lawyer	30%	10%	0%	50%
programmer	30%	40%	10%	50%
retiree	30%	30%	0%	50%
sales / marketing	30%	10%	0%	50%
scientist	30%	10%	0%	50%
self-employed	30%	30%	0%	50%
technician / engineer	30%	20%	0%	50%
tradesman / craftsman	30%	0%	0%	40%
unemployed	20%	10%	0%	20%
writer	30%	20%	0%	40%
other / not specified	0%	50%	0%	50%
Weighted Average	30%	37%	2%	49%

Table 8: Detailed results for using occupation ($\epsilon = 1.0$)

Job	Reduced-Attributes Method (Order)			Proposed Method
	1st	2nd	3rd	
academic / educator	30%	20%	0%	50%
artist	30%	20%	0%	50%
clerical / admin	30%	0%	0%	50%
college / grad student	30%	60%	0%	50%
customer service	30%	0%	0%	50%
doctor / health care	30%	10%	0%	40%
executive / managerial	30%	20%	0%	50%
farmer	20%	10%	0%	20%
homemaker	20%	0%	0%	40%
K-12 student	30%	0%	0%	50%
lawyer	30%	10%	0%	40%
programmer	30%	20%	10%	50%
retiree	30%	20%	0%	50%
sales / marketing	30%	0%	0%	50%
scientist	30%	0%	0%	40%
self-employed	30%	20%	0%	40%
technician / engineer	30%	0%	0%	50%
tradesman / craftsman	30%	0%	0%	30%
unemployed	10%	10%	0%	10%
writer	30%	0%	0%	30%
other / not specified	30%	20%	0%	50%
Weighted Average	30%	18%	1%	46%

Adventure & Sci-Fi” is (frequency of Action + frequency of Adventure + frequency of Sci-Fi) / 3. We additionally used as information the ratio $R_1 : R_2 : R_3 : \dots : R_{18}$, where R_i represents the number of purchases of i -th-order movies. Because the ratio information was derived from the original micro-data, its use would reveal private information. We therefore anonymized the ratio information before using it by adding noise to the information. However, the sensitivity of the ratio information was much smaller than that of values in the frequency table ³. Using the ratio information, the system recommended three first-order categories, five second-order categories, and two third-order categories.

4.2 Evaluation

We first used our recommendation system without anonymization and obtained the top ten recommendations for each user class. We used these recommendations as correct recommendations. We then used the system with anonymization. Three anonymization methods were used. The first one was a conventional method that added Laplacian noise with a sensitivity of $2^{18} - 1$. The second one was a reduced-attributes method that used up to i -th movie categories and ignored other categories. The second one was evaluated with $i = 1, 2, \text{ and } 3$. The third one was the proposed method. The recommendation accuracy (in percent) was measured in terms of the number of recommendations included in both the correct recommendations and the recommendations provided by the system. It was measured for each method and each parameter value (security parameter ϵ and order i).

Users were classified by gender and 7 age classes (i.e. 14 classes) in the first experiment, by occupation (21 classes) in the second experiment, and by gender, age, and occupation (294 classes) in the third experiment. For each experiment, two values of ϵ were used: 2.0 for weak anonymization and 1.0 for moderate anonymization.

Table 2 shows the results of the first experiment. The reduced-attributes method used up to first-, second-, and third-order movie categories. The conventional anonymization method could not produce any correct recommendations after anonymization. The accuracy of the reduced-attributes method was lower than that of the conventional method without anonymization because it used only some attributes. It had better accuracy than the conventional method after anonymization. The proposed method was better than the reduced-attributes method for most parameter values but worse than the second-order reduced-attributes method with weak anonymization ($\epsilon = 2.0$). Note that the advantage of the proposed method over the reduced-attributes method was larger when anonymization was stronger ($\epsilon = 1.0$). It means that the proposed method was robust against strong anonymization.

Tables 3 and 4 show the results of the second and third experiments. Again, the conventional method could not produce any correct recommendations after anonymization. The proposed method was better than the reduced-attributes method for all parameter values. The advantage of the proposed method over the reduced-attributes method was larger when anonymiza-

³The sensitivity of ratio information is 18.

tion was stronger ($\epsilon = 1.0$). The performance shown in Table 4 was worse than that shown in Tables 2 and 3. This is because the users were classified into too many classes, so the number of users (the number of purchase records) in each class was too small, resulting in a low signal-noise ratio.

Table 5 shows detailed results for the first experiment with $\epsilon = 2.0$. Accuracy is shown for each user class, anonymization method, and parameter value. The results of the conventional method are omitted because all values were 0%; i.e. none of the recommendations produced was correct. The proposed method was more accurate than the first- and third-order applications of the reduced-attributes method and slightly less accurate than the second-order application. Its accuracy was stable across user classes while that of the reduced-attributes method greatly depended on the user class.

Table 6 shows detailed results for the first experiment with $\epsilon = 1.0$. The advantage of the proposed method over the reduced-attributes method was larger than that with $\epsilon = 2.0$ (Table 5). Tables 7 and 8 show detailed results for the second experiment with $\epsilon = 2.0$ and $\epsilon = 1.0$, respectively. Like other experiments, the advantage of the proposed method over the reduced-attributes method was larger when $\epsilon = 1.0$ than when $\epsilon = 2.0$.

5 CONCLUSION

We have proposed a strategy for implementing differential privacy to reduce Laplacian noise and maintain the quality of anonymized statistics. A subset of attributes in the micro-data is used to generate statistics, reducing noise to add on the statistics with reduced attributes, and restoring information in the statistics by using relationships between the selected attributes and the other attributes. An example implementation (a recommendation system for movies using the MovieLens 1M dataset) demonstrated the viability of the proposed method.

REFERENCES

- [1] C. Dwork, Differential Privacy, ICALP, Vol. 4052, pp. 1–12 (2006).
- [2] L. Sweeney, K-anonymity: A Model for Protecting Privacy, IJUFKS, Vol. 10, No. 5, pp. 557–570 (2002).
- [3] A. Machanavajjhala, D. Kifer, J. Gehrke, and M. Venkatasubramanian, l-Diversity: Privacy Beyond k-Anonymity, TKDD, Vol. 1, No. 1 (2007).
- [4] L. Ninghui, L. Tiancheng, and V. Suresh, t-Closeness: Privacy Beyond k-Anonymity and l-Diversity, ICDE, pp. 106–115 (2007).
- [5] C. Skinner, Statistical Disclosure Control for Survey Data, Handbook of Statistics, Vol. 29, pp. 381–396 (2009).
- [6] A. Hundepool, J. Domingo-Ferrer, L. Franconi, S. Giessing, E. S. Nordholt, K. Spicer, and P.-P. de Wolf, Statistical Disclosure Control, A John Wiley & Sons (2012).
- [7] J. Soria-Comas and J. Domingo-Ferrer, Probabilistic k-Anonymity through Microaggregation and Data Swapping, ICFS, pp. 1–8 (2012).
- [8] C. C. Aggarwal and P. S. Yu, Privacy-Preserving Data Mining Models and Algorithms, Vol. 34, Springer (2008).
- [9] D. Kifer, B. Lin, Towards an Axiomatization of Statistical Privacy and Utility, PODS, pp. 147–158 (2010).
- [10] C. Dwork, F. McSherry, K. Nissim, and A. Smith, Calibrating Noise to Sensitivity in Private Data Analysis, TCC, pp. 265–284 (2006)
- [11] C. Dwork, Differential Privacy in New Settings, SODA, pp. 174–183 (2010).
- [12] C. Dwork, M. Naor, T. Pitassi, G. N. Rothblum, Differential Privacy under Continual Observation, STOC, pp. 715–724 (2010).
- [13] G. Kellaris, S. Papadopoulos, X. Xiao, D. Papadias, Differentially Private Event Sequences over Infinite Streams, VLDB Endowment, pp. 1155–1166 (2014).
- [14] C. Dwork, K. Kenthapadi, F. McSherry, I. Mironov, and M. Naor, Our Data, Ourselves: Privacy via Distributed Noise Generation, EUROCRYPT, Vol. 4004, pp. 486–503 (2006).
- [15] S. P. Kasiviswanathan and A. D. Smith, A Note on Differential Privacy: Defining Resistance to Arbitrary Side Information, IACR (2008).
- [16] N. Li, W. Qardaji, D. Su, On Sampling, Anonymization, and Differential Privacy or, K-anonymization Meets Differential Privacy, ASIACCS, pp. 32–33 (2012)
- [17] G. Cormode, M. Procopiuc, D. Srivastava, T. T. L. Tran Differentially Private Publication of Sparse Data, ICDDT (2011).
- [18] N. Shlomo, L. Antal, and M. Elliot, Measuring Disclosure Risk and Data Utility for Flexible Table Generators, JOS, Vol. 31, No. 2, pp. 305–324 (2015).
- [19] F. McSherry, I. Mironov, Differentially Private Recommender Systems, Building Privacy into the Net, SIGKDD, pp. 627–636 (2009).
- [20] S. Zhou, K. Ligett, and L. Wasserman, Differential Privacy with Compression, ISIT, pp. 2718–2722 (2009).
- [21] K. Chaudhuri, A. D. Sarwate, and K. Sinha, A Near-Optimal Algorithm for Differentially-Private Principal Components, JMLR, (2013).
- [22] C. Dwork, K. Talwar, A. Thakurta, and L. Zhang, Analyze Gauss: Optimal Bounds for Privacy-preserving Principal Component Analysis, STOC, pp. 11–20 (2014).
- [23] J. B. Schafer, J. A. Konstan, and J. Riedl, E-commerce Recommendation Applications, DMKD, Vol. 5, pp. 115–153 (2001).
- [24] D. K. Agarwal and B.-C. Chen, Statistical Methods for Recommender Systems, Cambridge Univ. Press (2016).
- [25] GroupLens, MovieLens 1M Dataset, available from <http://grouplens.org/datasets/movielens/1m> (2003).

A Classification Method of Unknown Malicious Websites Using Address Features of each Network Address Class

Shihori Kanazawa*, Yoshitaka Nakamura**, Hiroshi Inamura**, and Osamu Takahashi**

* Graduate School of Systems Information Science, Future University Hakodate, Japan

** School of Systems Information Science, Future University Hakodate, Japan
{g2116011, y-nakamr, inamura, osamu}@fun.ac.jp

Abstract - Recently, cyber-attacks through web sites such as Drive-by download attacks or phishing attacks increase rapidly. The attackers acquire personal information of users illegally by these attacks and economically damages to the users. The conventional detection method of malicious Web site uses a blacklist, and characteristics of the domain name. Since the domain name can be changed relatively easily, it is inappropriate for the method of detecting malicious Web sites to using domain names. In this paper, we propose a method of classifying Web sites as benign or malicious by using only a part of the network address, in order to reduce the classification cost. And, we evaluated the proposed method by cross-validation. As a result of evaluation, high classification accuracy was provided in IP address Class A, and we could confirm the effectiveness of the proposed distinction method.

Keywords: cyber-attack, Drive-by download, malicious Web site, network address, IP address of Class

1. INTRODUCTION

In recent years, the threat of attacks by viruses and malwares on the Internet are increasing year by year. Among them, attacks using Web sites are increasing rapidly. According to "2017 Information Security Ten major threats" announced by Information-technology Promotion Agency (IPA) in March 2017, the first place, the fourth place, and the sixth place are attacks against Web sites[1]. As an example of attack, cyber-attacks through Web sites such as Drive-by download attacks or phishing attacks increasing particularly rapidly. The attackers acquire personal information of users illegally by these attacks and inflicts economical damage[2]. Figure 1 shows the number of incidents occurring and damage amount data from Ref.[2] by National Police Agency. The number of the illegal acquisitions of the personal information that occurred by 2012 was 50 cases. However, the number of occurrences of the incidents has increased to 1400 cases until 2015. In order to prevent such damage, it is necessary to take measures to prevent users from accessing malicious Web sites.

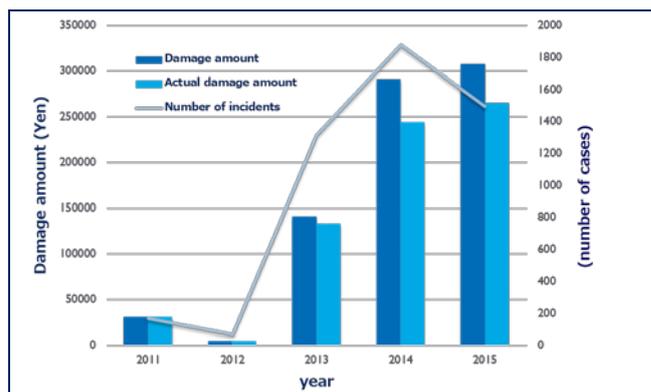


Figure 1: The number of incidents, damage amount, and actual damage amount

There are some methods to block access if the visited Web site is a malicious Web site to prevent users from accessing those sites. The first method is to use the Web reputation system[3]. The Web reputation system has a function of blocking malicious Web sites. If the domain name of the connection destination or the Web site is judged to malicious, the web reputation system blocks web access. In this way, the Web reputation system prevents damage caused by malicious programs and phishing. However, the Web reputation system can block access only to Web sites that have been confirmed to have carried out fraudulent acts such as virus distribution and phishing scams. The second method is to use Intrusion Prevention System. IPS supports sophisticated and advanced security threats such as bot attacks and DoS attacks that are considered to be difficult to protect only by general firewalls and anti-virus software. IPS examines the contents and behaviors of communication packets, and blocks web access if IPS detects communication that can be considered malicious. However, IPS can only detect known suspicious packets included in Web access communication.

Since the above two methods use known information such as information of suspicious packets included in known malicious Web sites, there is an advantage that the detection rate of known malicious Web sites is relatively high. However, these methods have drawbacks that can not be detected unknown malicious Web sites. Therefore, it is also unknown whether these methods can obtain sufficient accuracy. In order to solve such a problem, it is necessary to consider the detection conditions that enables detection including unknown malicious Web sites. Also, it is necessary to classify unknown Web sites into benign Web

sites and malicious Web sites. Therefore, we propose a method to detect and classify an unknown malicious Web sites.

This paper is constructed as follows. Section 2 mentions about researches related to our study. Section 3 mentions the requirements of the proposed method and our plan of the experiment. Section 4 mentions experimental results. Finally, this paper concludes, and discusses on future work in Section 5.

2. RELATED WORK

In recent years, some system has been developed to prevent the user from accessing to a malicious Web site. For example, one of them is Web reputation systems [3,4] for detecting a malicious Web site. This system utilizes a list of known malicious Web sites. The list of known malicious Web sites is called the blacklist. However, this system using the blacklist can not deal with unknown malicious Web sites.

There are some other researches to detect malicious Web sites including researches based on the features of URLs, researches based on the features of domain names, and researches based on the features of IP addresses. Each type of research is described in detail below.

First, we describe the research based on the feature of URLs. J. Ma et al., and K. Tanaka proposed a supervised learning approach for classifying URLs as normal or malicious based on the lexical structure of URLs [5]. This system can classify malicious domains and normal domains by the features that can be extracted from the DNS communication. Next, we describe the research based on the feature of domain names. I. Ryu et al. uses the features of the domain names as the detection condition [6]. A malicious domain name has a feature that the length of the domain name is 10 characters or more and alphanumeric characters are mixed. The malicious domain name has these features because it is often generated automatically using Fast-Flux attack method[7]. Fast-Flux uses computers infected with bots (botnet) to distribute viruses or guidance information for phishing attacks. As another feature, malicious domain names have many features that mix alphanumeric characters. L. Bilge et al. proposed a system that employs DNS analysis techniques to detect domains that are involved in malicious activity [8,9]. This system can classify malicious domains and normal domains by the features that can be extracted from the DNS communication. Malicious domain names are likely to be accessed by a client infected with malware. These methods are effective for detection of known Web sites, because these methods use the blacklist of domain names and URLs. However, these methods can not be easily maintained an up-to-date blacklist, because domain names can be easily and continuously generated or changed. As research based on IP addresses, Chiba et al. proposed a method of utilizing the feature of malicious IP addresses [10,11]. This research classifies malicious IP addresses and normal IP addresses by the feature of malicious IP addresses, because Cyber Attack is prone to use particular IP addresses[10,11,12]. However, this method has high cost because it is necessary to translate

an IP address into all bit strings.

From these studies, it can be said that the approach using blacklists tends to fail for versatile domain-based detection. And the approach of using domain names to classify malicious Web sites can be said to be difficult because domain name can be changed is easily. In this research, we propose a new method to detect a malicious Web site effectively by reducing avoidance from blacklist. This method uses only the domain name for detection in order to expand the detection range of the malicious Web sites. We also propose a method to classify Web sites at low cost by using a part of IP address in order to reduce the cost.

3. A METHOD FOR CLASSIFYING WEB SITES BY USING IP ADDRESS

3.1 Purpose of the study

In this paper, we propose a method to detect unknown malicious Web sites by increasing detection conditions using domain names of such Web sites. In the proposed method, the following two features of the malicious IP address are used to classify the Web site. It is known that a cyber-attack is prone to use particular IP addresses [10,11,12]. However, in order to use this IP address distribution feature, it is inefficient to classify malicious Web sites by comparing all the bit strings of IP addresses for each access communication. Figure 2 shows the usage distribution of malicious IP addresses.

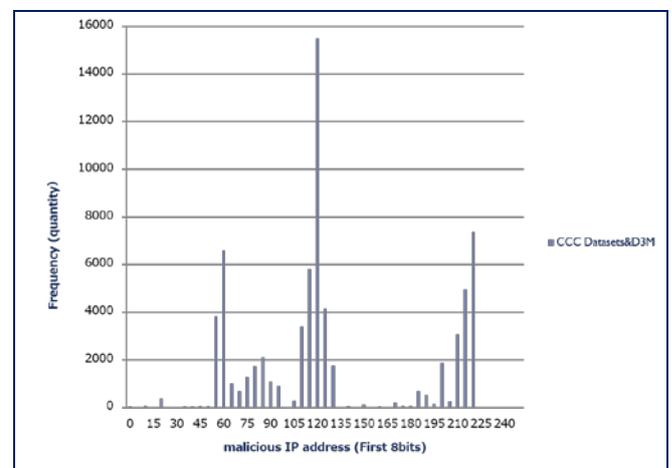


Figure 2: The usage distribution of malicious IP addresses

As shown in Figure 2, specific IP addresses are frequently used. Also, IP addresses can be classified into three IP address classes. Therefore, in this paper, we propose a method to reduce the cost of classifying benign or malicious Web sites using only the network address part according to the IP address class of each IP address.

3.2 Approach

Figure 3 shows the outline of the proposed system using the approach to lower classification cost.

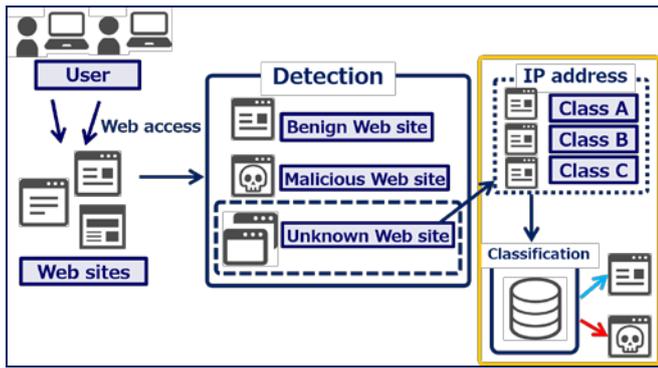


Figure 3: An approach to lowering the classification cost

Web sites accessed by clients can be classified into three categories: benign Web sites, malicious Web sites, and unknown Web sites. It is important how to deal with these unknown Web sites. The unknown Web sites are very likely to be malicious Web sites. In this approach, we will classify unknown Web sites to benign / malicious at low cost by using IP addresses of only unknown Web sites.

3.3 Proposed method

The proposed system consists of two methods. One is a method of detecting unknown Web sites by removing known malicious and benign Web sites. The other method is to classify whether an unknown Web site is malignant or benign. Details of the detection method are described in Section 3.4, and details of the classification method are explained in Section 3.5. Figure 4 shows the overview of the whole proposed system.

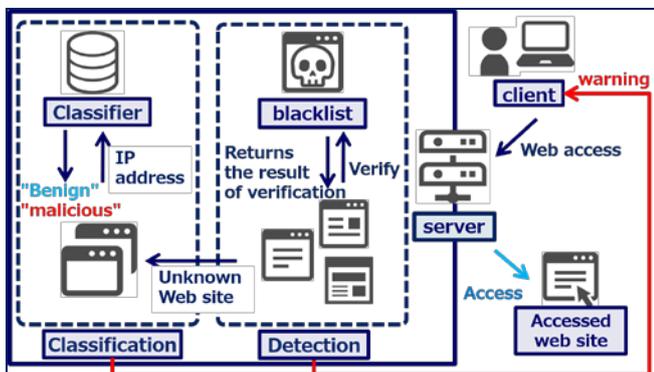


Figure 4: The overview of the proposed method

The proposed system consists of a detection unit and a classification unit. In the detection unit, unknown malicious Web sites that do not exist in the known blacklist are taken out. First, when the DNS server is accessed by the client, the detection unit checks the accessed Web site against the blacklist of known malicious Web Sites. If the Web site is detected to be unknown Web sites, the detection unit sends the IP address of the Web site to the classification unit to use for classification. The classification unit uses the features of this IP address to classify whether the Web site is benign or malicious. If the classification result is benign, this system allows the client to access the Web site. If the classification result is malicious, this system alerts the client,

and update the blacklist in order to keep the classifier up to date.

3.4 The method of detection unknown Web sites

Figure 5 shows details of the detection unit which detect unknown Web sites.

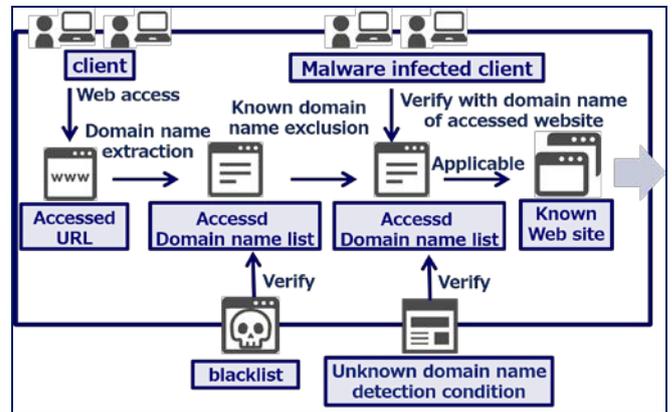


Figure 5: Details of detection method of unknown Web site

This detection unit uses domain names to detect unknown Web sites. The domain names are obtained from Web site URLs. The domain name is checked against the blacklist in order to exclude the known malicious domain. If the domain name of the Web site is not on the blacklist, the domain name is checked against the detection conditions based on the domain name. Unknown domain name detection condition has features of known malicious domain name. Also, in order to select an unknown Web site, the proposed system uses the domain name accessed by the client infected with malware (malware infected client).

3.5 The method of classifying malicious Web sites

The classification unit of the proposed system consists of two phases. The phase 1 is the construction of a classifier for generation of feature vectors using the training datasets (3.5.1). The phase 2 is the classification of the test dataset using constructed classifier (3.5.2). The training datasets are consisted by known malicious and benign Web sites. Test datasets are unknown Web sites.

3.5.1 Construction of classifier using training datasets

It is necessary for construction of classifier to generate feature vectors. Feature vectors have various features of the training datasets. And the number of feature vectors is referred to as the dimension number. In this paper, we propose a method of classifying Web sites as benign or malicious by using a low number of dimensions, in order to reduce the classification cost. Figure 7 shows the method to generate feature vectors.

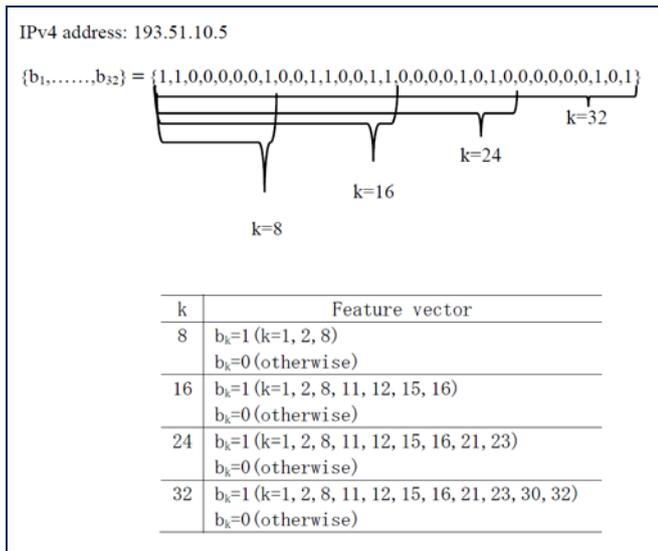


Figure 6: The example of generating feature vectors

First, the IP address of the training datasets is translated into bit strings. Secondly, all bit strings are represented as k-dimensional vector $\{b_1, \dots, b_k\}$. Three types of feature vectors are generated according to the IP address class such as Class A, Class B, and Class C. A vector of 8 dimensions in the case of Class A, a vector of 16 dimensions in the case of Class B, and a vector of 24 dimensions in the case of Class C is generated. Finally, generated feature vectors of malicious are labeled “1”, and generated feature vectors of benign are labeled “0”. Table 1 shows that an example of labeling feature vectors of the training datasets.

Table 1: Example of labeling feature vectors of training datasets

IP address	Feature vector	Label
193.51.10.5	1,1,0,0,0,0,0,1,0,0,1,1,0,0,1,1	1
10.10.10.10	0,0,0,0,1,0,1,0,0,0,0,0,0,1,0,1	1
203.4.12.89	1,1,0,0,1,0,1,1,0,0,0,0,0,1,0,0	0
...

The classifier used in this paper is a Support Vector Machine (SVM), which is one of pattern identification methods. Reference[5] written by J. Ma et al. clearly indicated that malicious Web sites are detected with high accuracy by using SVM. The proposed system constructs three classifiers based on feature vectors described above.

3.5.2 Classification of the test datasets using constructed classifier

The test datasets are classified by the constructed classifiers in the phase 1. Figure 8 shows that classification has three steps.

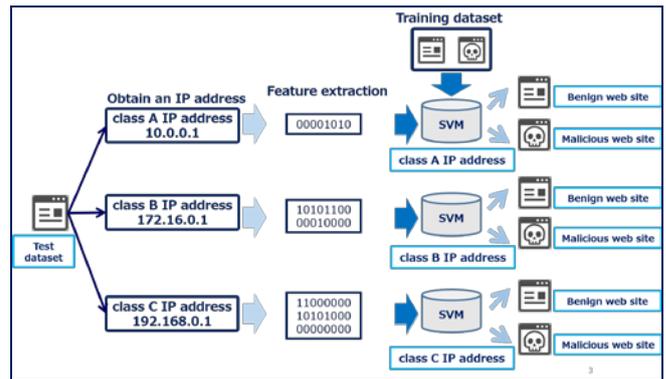


Figure 7: The process of the classification

First, the classification section obtains the IP address from the communication log of the test data set. A feature vector is generated from this obtained IP address. Finally, the feature vectors of the test dataset are classified as benign or malicious by the classifier constructed in the phase 1.

3.6 EVALUATION

The aim of the proposed system is to classify unknown Web sites as malicious or benign. The system may show that IP address of Class A achieves high accuracy with the first 8 bits. IP address of Class B shows high accuracy with the first 16 bits. IP address of Class C shows high accuracy with the first 24 bits.

We conduct an experiment to confirm the effectiveness of the classifier. We evaluated the constructed classifier at three points of accuracy, precision, recall rate. These three points are called evaluation indices.

This paper defines the correct classification of an actual malicious IP address into a malicious category as a true positive (TP), the incorrect classification of actual benign IP addresses into a malicious category as a false positive (FP), the incorrect classification of actual malicious IP addresses into a benign category as false negative (FN), and the correct classification of actual benign IP address into a benign category as a true negative (TN). Accuracy (1), precision rate (2) and recall rate (3) are calculated as follows:

$$Accuracy = (TP + TN)/(TP + TN + FP + FN) \dots(1)$$

$$Precision = TP/(TP + FP) \dots(2)$$

$$Recall = TP/(TP + FN) \dots(3)$$

The dataset of benign and malicious are obtained from Malware Workshop Datasets [13]. The dataset is created with the ratio of malicious IP addresses and benign IP addresses of 8:2, 5:5, 2:8. The malicious training datasets was created from CCC datasets (2008 - 2011) and D3M (2010 - 2015). The benign training datasets were created from Alexa's top sites 50,000 (2016) [14] and NCD in MWS Cup (2014). IP addresses of Class A have 49164 malicious IP addresses and 40667 benign IP addresses. IP addresses of Class B have 3523 malicious IP addresses and 10735 benign IP addresses. IP addresses of Class C have 75000 malicious IP addresses and 14288 benign IP addresses. The

classification experiments are conducted with 5-way cross validation.

4. Result

We carry out the experimental evaluation in four ways as defined in Figure 7. We defined the classification using first 8 bits for Case 1, first 16 bits for Case 2, first 24 bits for Case 3, first 32 bits for Case 4. Table 2 shows the experimental result of classifying the IP address belonging to the IP address of Class A.

Table 2: Experimental result of IP address of Class A

	Accuracy	Precision	Recall
Case1(k=8)	84.06079	89.86656	90.25327
Case2(k=16)	83.74358	89.79705	89.89365
Case3(k=24)	83.76079	89.8188	89.89058
Case4(k=32)	83.87882	89.95386	89.8875

The recall rate achieved the highest value in Case 1. Accuracy and Recall are the highest in Case 1, while Precision is Case 4 the highest.

Table 3 shows the experimental result of classifying the IP address belonging to the IP address of Class B.

Table 3: Experimental result of IP address of Class B

	Accuracy	Precision	Recall
Case1(k=8)	83.54143	87.85521	92.15756
Case2(k=16)	81.69694	88.16298	89.07026
Case3(k=24)	82.94552	88.61024	90.27679
Case4(k=32)	83.25766	88.76131	90.5252

The recall rate achieved the highest value in Case 1. The accuracy also achieved the highest value at Case 1. The precision rate achieved the highest value at Case 4. The accuracy, precision rate and recall rate of Case 2 were lower than those of Case 1, Case 3, and Case 4.

Table 4 shows the experimental result of classifying the IP address belonging to IP address of Class C.

Table 4: Experimental result of IP address of Class C

	Accuracy	Precision	Recall
Case1(k=8)	81.78191	87.19137	90.52493
Case2(k=16)	81.20101	86.9819	89.965
Case3(k=24)	81.243	87.00101	90.0
Case4(k=32)	81.222	86.56467	90.58618

The precision rate achieved the highest value in Case 1. The accuracy also achieved the highest value in Case 1. The recall rate achieved the highest value in Case 4. The accuracy, precision rate and recall rate of Case 3 were lower than those of Case 1, Case 2, and Case 4.

4.2 Discussion

The IP address belonging to Class A IP address seems to obtain the best result because the IP address frequently used for malicious activities and the IP address not used is clearly classified. In addition, Class A is considered to have sufficient information suitable for classification of a malicious Web site because it has a large number of distributions to users. Therefore, classification using network address is effective on Class A IP address. The Class B IP address is not used for many malicious activities, so few features is found in these addresses. Also, in the case of the Class C IP address, there is a possibility that the specific IP address has lowered three evaluation indices. In addition, because Class C is easiest to distribute to users among all IP address classes, malicious IP address is likely to be easily changed. Therefore, it is necessary to analyze the IP address which reduces evaluation indices of classification by IP address of Class.

We compared the usage status of IP addresses that were identified by mistake and IP addresses used for malicious activities. Figure 9 shows the result of comparing the usage status of IP addresses used for malicious activities with the IP addresses belonging to Class B IP address that were incorrectly classified.

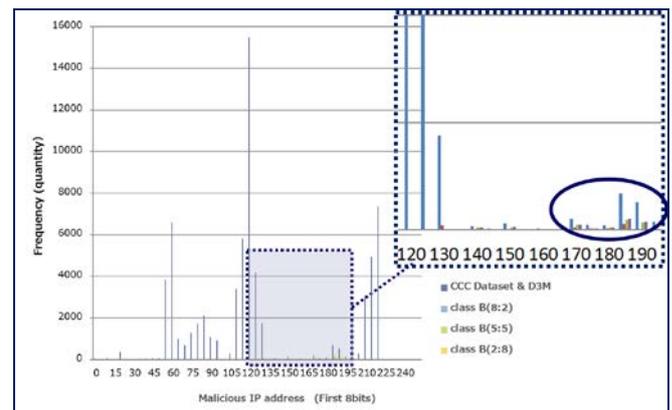


Figure 8: The result of comparing the usage status of IP addresses (Class B)

The IP address whose leading 8 bits are around 120 is often used for malicious activities, so the features of malicious activities are considered to exist. However, the incorrectly classified IP addresses tend to be concentrated on areas where the features of malicious activities do not exist.

Figure 10 shows the result of comparing the usage status of IP addresses used for malicious activity with the IP addresses belonging to Class C IP address that were incorrectly classified.

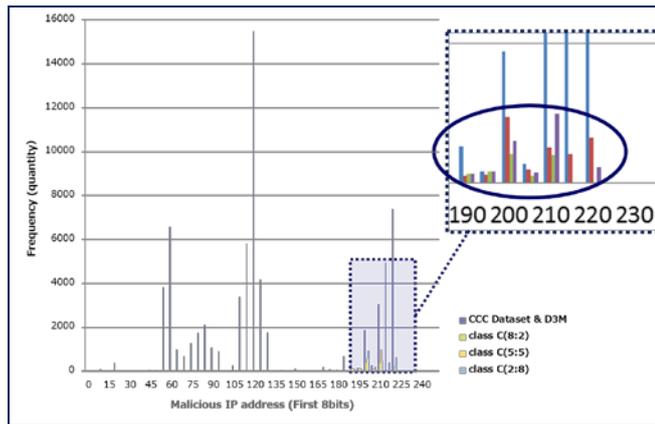


Figure 9: The result of comparing the usage status of IP addresses (Class C)

Since IP addresses belonging to Class C IP address are partially used for malicious activities, it is considered that there are the features of malicious activities. However, the IP address incorrectly classified tends to be concentrated on area where there are the features of malicious activities do not exist.

4.3 Limitation

There may be some limitations in this research. The evaluation indices of Class B and Class C IP address is low in the current training datasets. It is probably due to the amount of Class B and Class C IP address in the datasets. The number of IP addresses belonging to Class B IP address is very small. On the other hand, the number of IP addresses belonging to Class C IP address is very large. Therefore, it is necessary to devise selection of the IP address of the training dataset.

4.4 Future Work

Further research should be done to classify malicious Web sites. First, it is necessary to carefully select IP addresses of training datasets belonging to Class C and Class B IP address. The evaluation indices may be improved by this selection. Secondly, it is necessary to analyze trends of datasets by year. The features of Class C may change from year to year. Thirdly, it is necessary to consider the way to return to the benign IP address in the method classifying by IP addresses. Malicious IP addresses are used for malicious activities for a while. However, when malicious IP addresses are used frequently, malicious activities become to use the other IP addresses for avoiding detection. The IP address that is no longer used may be used by someone else. Finally, it is necessary to confirm that the efficiency of the classification is improved by the proposed system.

5. CONCLUSION

In this paper, we propose a method of detecting unknown Web sites and classifying Web sites as benign or malicious

by using only a part of the network address, in order to reduce the cost of the classification. And, we evaluate the accuracy of the proposed classification for each IP address Class by cross-validation. As a result of evaluation, high classification precision was provided in Class A IP address. From this result, we confirm the effectiveness of the proposed classification method. But, there may be some limitations in this system. First, the evaluation indices of Class B and Class C IP address are relatively low in the current training datasets. Secondly, the IP addresses classified as malicious cannot return to benign IP addresses.

In the future, further research should be done to investigate the feature of malicious Web sites. First, it is necessary to carefully select IP addresses of the training datasets belonging to Class B and Class C IP address. This selection of IP addresses may improve the evaluation indices. Secondly, it is necessary to analyze trends of datasets by year. The features of Class C IP address may change from year to year. Thirdly, it is necessary to consider the way to return to the benign IP address in the method classifying by IP addresses. Finally, it is necessary to confirm that the efficiency is improved by the proposed system.

REFERENCES

- [1] Information-technology Promotion Agency, Japan, "10 Major Security Threats" <<https://www.ipa.go.jp/files/000058504.pdf>> [Accessed May 31, 2017] (in Japanese).
- [2] National Police Agency, "Public information of the National Police Agency: About the occurrence situation of illegal remittance offenses related to Internet banking in Heisei 26," <https://www.npa.go.jp/cyber/pdf/H270212_banking.pdf> [Accessed January 17, 2017] (in Japanese).
- [3] TREND MICRO Incorporated, "Web reputation," <<http://www.trendmicro.co.jp/why-trendmicro/spn/features/web/index.html>> [Accessed May 31, 2017] (in Japanese).
- [4] M. A. Rajab, L. Ballard, N. Jagpal, P. Mavrommatis, D. Nojiri, N. Provos, and L. Schmidt, "Trends in circumventing web-malware detection," Google, Google Technical Report, 2011.
- [5] J. Ma, L. K. Saul, S. Savage, and G. M. Voelker, "Beyond blacklists: learning to detect malicious web sites from suspicious urls," Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD'09), pp. 1245–1254, 2009.
- [6] I. Ryu, "Detection method of malicious site by DNS information," Master's thesis of Waseda University 2012.
- [7] Hitachi Solutions, Ltd. "Information security blog," <<http://securityblog.jp/words/2898.html>> [Accessed January 17, 2017] (in Japanese).
- [8] L. Bilge, E. Kirde, C. Kruegel, and M. Balduzzi, "Exposure Finding Malicious Domains Using Passive DNS Analysis," Proceedings of the 18th Annual

Network & Distributed System Security Symposium (NDSS Symposium 2011), 2011.

- [9] K. Tanaka, A. Nagao, and M. Morii, "Extracting Malicious Website from DNS Log-Analysis Method and Anonymity-," Proceedings of the Computer Security Symposium 2013(CSS2013), pp.132-138, 2013 (*in Japanese*).
- [10] D. Chiba, K. Tobe, T. Mori, and S. Goto, "Detecting Malicious Websites by Learning IP Address Features," Proceedings of the IEEE/IPSJ 12th International Symposium on Applications and the Internet(SAINT2012), pp.29-39, 2012.
- [11] D. Chiba, T. Mori, and S. Goto, "Deciding priority crawling in searching for malicious websites," Proceedings of the Computer Security Symposium 2012(CSS2012), pp.805-812, 2012 (*in Japanese*).
- [12] D. Chiba, T. Yagi, M. Akiyama, and T.Mori, "Correlation Analysis Between IP Addresses Used in Variety of Attacks," Proceedings of the Computer Security Symposium 2011(CSS2011), pp.185-190, 2013 (*in Japanese*).
- [13] M. Kamizono, M. Akiyama, T. Kasama, J. Murakami, M. Hatada, and M. Terada, "Datasets for Anti-Malware Research ~MWS Datasets 2015~," he Special Interest Group Technical Reports of IPSJ Vol.2015-CSEC-70, No.6, pp. 1-8, 2015. (*in Japanese*).
- [14] Alexa Internet, Inc., "The top 500 sites on the web," <<http://www.alexa.com/topsites>> [Accessed January 17, 2017].

Online Magnitude Fluctuation Analysis for Anomaly Detection in Equipment Condition Monitoring

Makoto Imamura^{*}, Daniel Nikovski^{**}, Amir-massoud Farahmand^{**}, Harumi Watanabe^{*}

^{*} School of Information and Telecommunication Engineering, Tokai University, Japan

^{**} Mitsubishi Electric Research Laboratories, USA

imamura@tsc.u-tokai.ac.jp

Abstract - In industrial domains, equipment condition monitoring (ECM) has attracted much attention recently, especially as the Internet of Things (IoT) has been emerging and advancing rapidly. Fluctuation of time series generated by industrial equipment has been recognized as an important problem, and a variety of methods have been developed for its analysis. For example, Fourier analysis is a standard method for diagnosing machine failure from stationary time series such as the vibration of a motor, and cycle counting methods are often used for diagnosing fatigue damage from unsteady and random time series such as stress-strain cycle. This paper proposes a new magnitude fluctuation feature for unsteady and random data, and also describes an efficient online algorithm for computing it. Magnitude fluctuation is defined by a convex pattern which consists of an upward trend leg and a downward trend leg, and it generalizes cycle counting methods in the sense that it captures not only the amplitude of fluctuation, but also its duration. This paper also describes how our proposed method is applicable to anomaly detection for ECM by detecting the changes in the appearance frequency and the amplitude of the convex patterns in sensor time series.

Keywords: Magnitude Fluctuation Analysis, Anomaly Detection, Feature Extraction, Time Series Datamining, Equipment Condition Monitoring, Online Algorithm

1 INTRODUCTION

As the Internet of Things (IOT) [1] has been emerging and growing, sensor big data that is streamed from various equipment in power plant, industrial facilities, and buildings can be available for monitoring, diagnosis, energy-saving, productivity improvement, quality management and marketing. As a result, industry has paid much attention to the use of big sensor data generated from equipment or facility in order to create smart society.

Equipment Condition Monitoring (ECM) is a typical service using sensor big data, and datamining techniques are key components to make ECM smarter [2]. This paper proposes a new magnitude fluctuation feature for unsteady and random data as a datamining technique, and also describes an efficient online algorithm for sensor big data.

After mechanical equipment is operated for a long time, convex-shaped spikes can often be observed in the torque current of motors, or in the pressure inside a pipe, because of frictional wear, adhesion of foreign substances, etc. We propose a method for extracting convex-shaped spikes in transient sensor data for detecting anomaly or degradation of

equipment (Fig. 1). Nevertheless, sensor data include various convex-shaped patterns, such as controlled operating patterns, random noise, and symptoms of degradation. We also propose a new feature to extract a convex curve that is a symptom of degradation by using the height of the convex curve (its amplitude) and the duration of time needed to form the convex curve (its length).

Our proposed method has the following two characteristics:

(1) It is parameter free

Our proposed new feature is parameter free. Most of time series analysis methods need parameter tuning, which is one of the main difficulties for data analysts. One such typical parameter is window size. For example, AR models and then discord algorithms, which are popular time series analysis methods, need the correct window size to be specified.

(2) It allows for online processing

Anomaly detection algorithms for equipment condition monitoring often need to operate online on streaming data, and must be fast enough for real-time processing. Our algorithm to calculate the proposed feature is online, and its computational complexity is $O(n)$, where n is the length of the time series. This makes it suitable for condition monitoring.

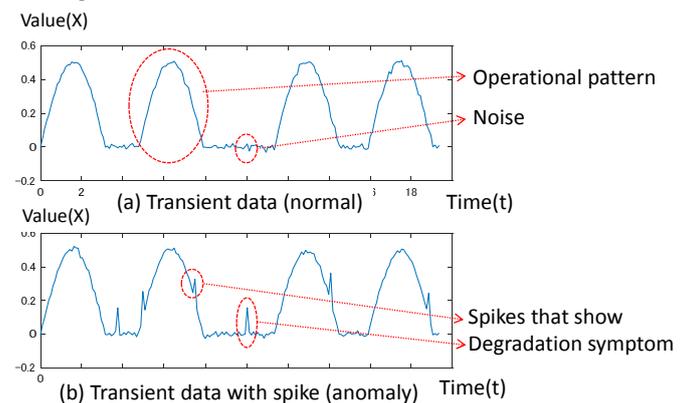


Figure 1: Transient data with spikes

Magnitude fluctuation for unsteady data has been studied from data mining point of view by Fink et al [3]. They proposed a concept of leg and its search method to find a global trend in a time-series including small variations such as noise. The dotted lines in Fig. 2 are examples of legs. Both lines show the global upward trend that includes local up-down segments. However, their method treats only single legs so that it can find an upward or downward trend, but can't catch the magnitude of fluctuations. We developed a

leg vibration analysis that can calculate the frequency of fluctuations in time-series that includes upward trends and downward trends that can appear alternately and iteratively [4][5]. These previous works define a leg frequency for a given window size and an amplitude, therefore it needs a window size and an amplitude as parameters. On the other hand, this paper defines the amplitude of the maximal convex curve at each time of a given time series as a new feature. The amplitude needs no parameter so this feature is parameter free.

Furthermore we propose an online algorithm to calculate the amplitude of the maximal convex curve as well as the previous works on the leg frequency analysis. With regards to computational complexity, that of the previous leg frequency calculation algorithm is $O(nw)$, but that of amplitude calculation algorithm can be reduced to $O(n)$ to reflect the fact that a window size is not a parameter.

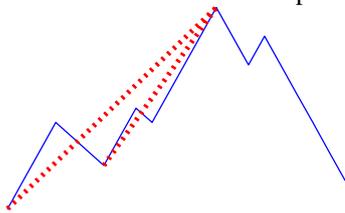


Figure 2: Leg

2 MAXIMAL CONVEX CURVE

This section defined the amplitude of the maximal convex curve at each time of a given time series as a feature which shows the degree of magnitude fluctuation of the time series. This feature depends only on time and the time series, so the analysis method based on this feature is parameter free. That is, it is needless to tune parameters when we use this feature. On the other hand, most of the features of time series depend on at least window size, so how to select a window size is often a problem.

The maximal convex curve at each time t is defined as a pair of the maximal leg from t towards left and that toward right. But naive definition of a maximal leg makes the amplitude unstable. Therefore, we introduce an extended maximal leg from t towards left or right to obtain a robust definition.

Definition: time series X , subsequences $X[p:q]$

A Time Series $X=[x_1, \dots, x_m]$ is a continuous sequence of real values. The value of the i -th time point is denoted by $X[i] = x_i$.

A subsequence $S = [x_p, x_{p+1}, \dots, x_q] = X[p:q]$ is a continuous subsequence of X starting at position p and ending at position q . We denote the length of a subsequence S by $len(S) \equiv q - p + 1$

Definition: Leg

Let X be a time series. We define a leg by a subsequence $L = X[l:r]$ that satisfies the conditions below.

$$\forall i. l < i < r \quad (X[r] - X[i])(X[i] - X[l]) > 0$$

That is, a subsequence $X[l:r]$ has a maximum and a minimum at the terminal points l, r .

If $X[r] - X[l] > 0$, a leg L is called an upward leg.

If $X[r] - X[l] < 0$, a leg L is called a downward leg.

Fig 3. shows the examples of upward and downward legs.

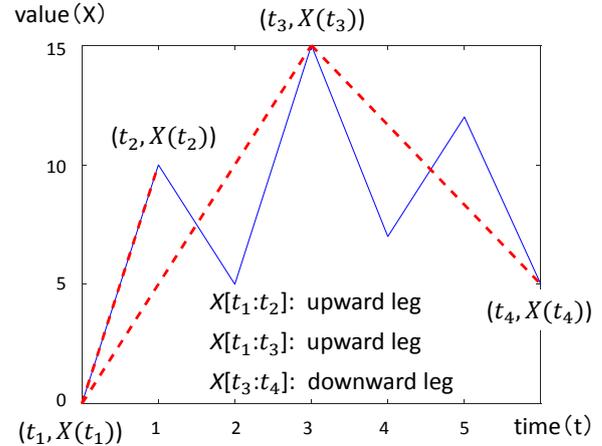


Figure 3: Leg

Definition: Sign and amplitude of a leg

We define the *sign* and *amplitude* of a leg $L = X[l:r]$ by the functions below. We denote them by *amp* and *sign* respectively:

$$amp(L) = abs(X[r] - X[l]).$$

The absolute function $abs(a)$ means the absolute value of a .

$$sign(L) = 1, \text{ if } (X[r] - X[l]) > 0$$

$$= 0, \text{ if } (X[r] - X[l]) = 0$$

$$= -1, \text{ if } (X[r] - X[l]) < 0$$

By the above definition, the sign of an upward leg is plus and the sign of a downward leg is minus.

Definition: A Maximal Leg from t toward left

Let X be a time series and t be a time point in X .

We define a *Maximal Leg from t toward left* by a leg $L = X[l:t]$ that satisfies the following condition.

For any $l' < l$, $X[l':t]$ is not a leg the amplitude of which is larger than that of $X[l:t]$.

That is,

for any l' such that $l' < l$ and

$$sign(L)(X[t] - X[l]) \leq sign(L)(X[t] - X[l'])$$

some j such that $l' \leq j < l$ exists and j satisfies

$$sign(L)(X[t] - X[j]) \leq 0$$

Value(X) $X[t_3:t_1]$: maximal leg from t_1 toward left

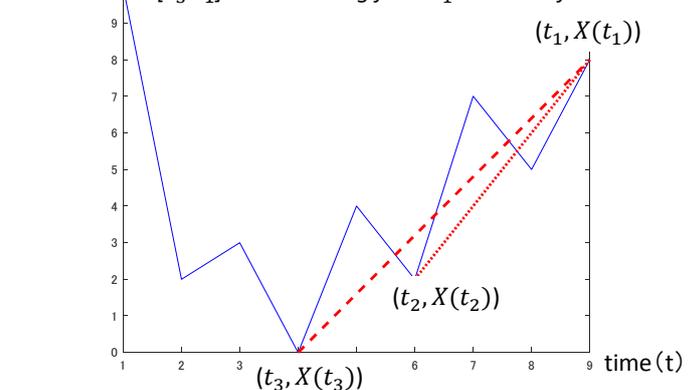


Figure 4: Maximal leg from t toward left

Fig 4. shows the example of a maximal leg from t toward left. $X[t_3:t_1]$ is a maximal leg from t toward left. On the other hand, $X[t_2:t_1]$ is a leg but not maximal leg from t toward left.

Definition: A Maximal Leg from t toward right

We define a *Maximal Leg from t toward right* by a leg $L = X[t:r]$ that satisfies the following the same as "Maximal Leg from t toward left":

For any $r < r'$, $X[t:r']$ is not a leg the amplitude of which is larger than that of $X[t:r]$.

Definition: Convex curve at t in X

Let X be a time series and t be a time point in X .

We define a *convex curve at t in X* by a subsequence $S = X[l:r]$ that satisfies the following conditions.

- (i) $l < t < r$
- (ii) $X[l:t]$ is a leg.
- (iii) $X[t:r]$ is a leg.
- (iv) $\text{sign}(X[l:t]) \text{sign}(X[t:r]) < 0$

We denote it by $X[l:t:r]$, and call t the *vertex* of the convex curve. l and r are called *left terminal* and *right terminal* of the vertex, respectively. We call an interval $[l:t]$ *support* of the convex curve.

Definition: Maximal Convex Curve at t in X

We define a *maximal convex curve at t in X* by a subsequence $S = X[l:r]$ that satisfies the following conditions.

- (i) $l < t < r$
- (ii) $X[l:t]$ is a maximal leg from t toward left.
- (iii) $X[t:r]$ is a maximal leg from t toward right.

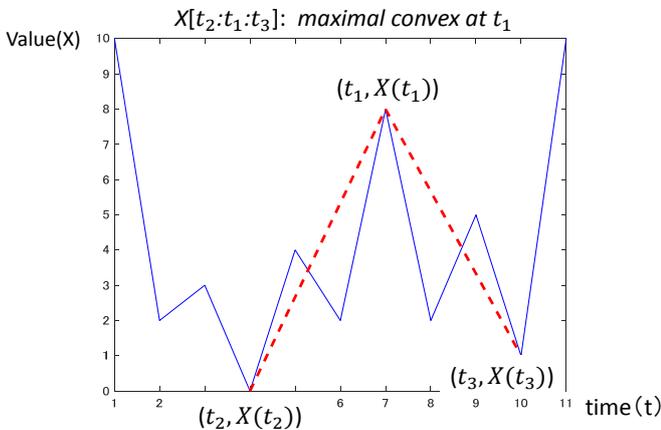


Figure 5: Maximal convex curve

Fig 5. shows the example of a maximal convex curve. $X[t_2:t_1:t_3]$ is a maximal convex curve at t_1 .

Definition: Signed Amplitude of a maximal convex curve

Let $C = X[l:t:r]$ be a maximal convex curve.

We define the *amplitude* $\text{amp}_c(X, t)$, *sign* $\text{sign}_c(X, t)$ and *signed amplitude* $\text{signedAmp}_c(X, t)$ of a maximal convex curve at t in X , respectively, by the functions below:

$$\text{amp}_c(X, t) \equiv \min(\text{amp}(X[l:t]), \text{amp}(X[t:r]))$$

$$\text{sign}_c(X, t) \equiv \text{sign}(X[l:t])$$

$$\text{signedAmp}_c(X, t) \equiv \text{sign}_c(X, t) \times \text{amp}_c(X, t)$$

If t is not a vertex of a maximal convex curve, we define the amplitude and sign at t to be zero.

The above definition of a maximal condition is not robust in the sense that even a small change of vertex value can make a large change of amplitude value. For example, in Fig. 6, suppose that $X(t_2) = X(t_4)$ and $X(t_1) = X(t_5)$. If $t_2 = t_3 = t_4$, a maximal convex curve at t_2 is $X(t_1:t_2:t_5)$ and its amplitude is $(X(t_2) - X(t_1))$. If $X(t_3)$ is just a little less than $X(t_2)$, a maximal convex curve at t_2 is $X(t_1:t_2:t_3)$, and a maximal convex curve at t_4 is $X(t_3:t_4:t_5)$. And the amplitude of each convex curve is $(X(t_2) - X(t_3))$. That is, just a small change of X is capable of making a great change of the amplitude.

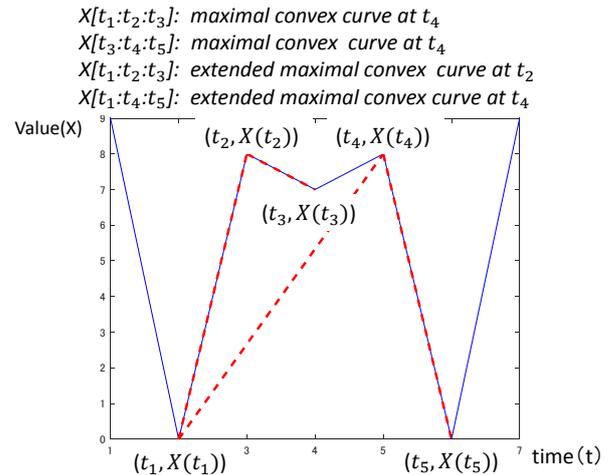


Figure 6: Extended convex curve

In real sensor data, even real valued data may have discrete values by the limitation of a sensor or a measuring instrument, such that they often might have the same values. Therefore, the above is not an atypical contrived case.

We introduce the concept of an extended leg in order to obtain a robust definition of a maximal convex curve.

$X[t_2:t_4]$ is an Extended leg
 $X[t_1:t_4]$ is not an Extended leg
 $X[t_1:t_3]$ is not an Extended leg

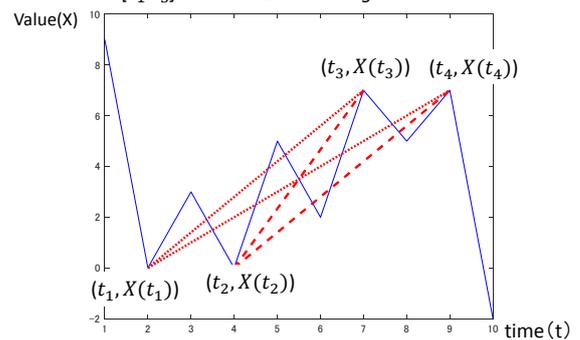


Figure 7: Extended leg

Definition: Extended Leg

We define an *Extended Leg* by the sequence $L = X[l:r]$ that satisfies the condition below.

$$\forall i. l < i < r \quad (X[r] - X[i])(X[i] - X[l]) > 0 \\ \vee X[i] = X[r]$$

Fig 7. shows the example of an extended leg: $X[t_2:t_4]$ is an extended leg. On the other hand, $X[t_1:t_4]$ and $X[t_1:t_3]$ are not extended legs.

"Maximal Extended Leg from t toward left and right" can be defined by the same way as "a Maximal Leg from t toward left and right", respectively.

Definition: A Maximal Extended Convex at t in X

Let X be a time series and t be a time point in X .

We define a maximal extended convex curve at t in X by a subsequence $C = X[l:r]$ that satisfies the following conditions. We denote it by $X[l:t:r]$

- (i) $l < t < r$
- (ii) $X[l:t]$ is a maximal extended leg from t toward left.
- (iii) $X[t:r]$ is a maximal leg from t toward right.

Please note that the left side is extended but the right side is not extended in the above definition, so that we do not count the larger amplitude twice, as shown in Fig. 5.

Fig. 6 shows the example of extended convex curves. $X[t_1:t_2:t_3]$ and $X[t_1:t_4:t_5]$ are extended maximal convex curves. On the other hand, $X[t_1:t_2:t_5]$ is not an extended maximal convex curve.

Definition: A maximal balanced convex curve at t in X

Let a subsequence $C = X[l:t:r]$ be a maximal convex curve at t . A convex curve the amplitude of which is the same as that of C is called a maximal balanced convex curve at t .

In other words,

- if $\text{amp}(X[l:t]) > \text{amp}(X[t:r])$
 $X[l_b:t:r]$ where l_b is the largest l' which satisfies
 $\text{amp}(X[l':t]) \geq \text{amp}(X[t:r])$
and $l \leq l'$.
- else if $\text{amp}(X[l:t]) < \text{amp}(X[t:r])$
 $X[l:t:r_b]$ where r_b is the smallest r' which satisfies
 $\text{amp}(X[t:r']) \geq \text{amp}(X[l:t])$
and $r' \leq r$.
- else $X[l:t:r]$.

We define the *amplitude*, *sign* and *signed amplitude* of a maximal balanced convex curve at t as with those of a maximal convex curve. We also define the *length* $\text{len}_c(X, t)$ of a maximal balanced convex at t curve by the function below:

Let a subsequence $C = X[l_t:t:r_t]$ be a balanced maximal convex curve at t .

$$\text{len}_c(X, t) \equiv r_t - l_t + 1$$

If t is not a vertex of a maximal convex curve, we define $\text{len}_c(X, t)$ to be 0.

3 ONLINE ALGORITHM

This section shows an online algorithm to calculate the signed amplitude and the length of a maximal convex curve in time series. The computational complexity of a naive algorithm derived by the definition of a maximal convex curve is $O(n^2)$, but we will proposed a fast algorithm whose computational complexity is $O(n)$ based on the results in the preceding section 2.2.

The obtained maximal convex at a given time does not depend on the order of the application of a middle leg reduction and a terminal leg reduction by the propositions. Therefore, we can get an online algorithm by the repetition of executable reductions while getting the value from the beginning.

Fig 8. shows an algorithm that computes the signed amplitude and the length of maximal balanced convex curves. In Fig. 8, X is an input time series which is implemented as a one-dimensional array and A and L are the output signed amplitude and length, respectively. Both are implemented as one-dimensional arrays.

Algorithm: maximalConvexAmplitude (X)	
Input:	X : time series
Global Var.:	S :Stack, P :Previous value, M :Previous local maximum
Output:	A :amplitude, L : length
01	// (0) Initialization
02	$A := \text{zeros}(\text{len}(X));$ // All the value of A is zero.
03	$L := \text{zeros}(\text{len}(X));$ // All the value of L is zero.
04	$S = [];$ // Stack
05	$P := \text{undef};$ // Previous time used in <i>getNextLocalMaximum</i>
06	$M := \text{undef};$ // Previous local maximal time
07	// Main loop
08	for $i := 1$ to $\text{len}(X)$; // $\text{len}(X)$ is the length of X
09	// (1) Prop.1: Local maximal transformation
10	$i := \text{getNextLocalMax}(X, i)$
11	$S = \text{push}(S, X[i]);$
12	// (2) Prop. 2: Middle leg reduction
13	if $\text{len}(S) \geq 4$
14	if $\text{abs}(X(S[3]) - X(S[4])) > \text{abs}(X(S[2]) - X(S[3]))$
15	and $\text{abs}(X(S[1]) - X(S[2])) \geq \text{abs}(X(S[2]) - X(S[3]))$;
16	$A(S[3]) = X(S[3]) - X(S[2]);$
17	$L(S[3]) = S[2] - \text{find_left}(X, S[3], S[4]) + 1;$
18	$A(S[2]) = X(S[2]) - X(S[3]);$
19	$L(S[2]) = \text{find_right}(X, S[2], S[1]) - S[3] + 1$
20	$S = \text{pop}(S, [2,3]);$
21	end if
22	// (3) Prop. 3: Terminal leg reduction
23	if $\text{len}(S) == 3$;
24	if $\text{abs}(X(S[1]) - X(S[2])) \geq \text{abs}(X(S[2]) - X(S[3]))$;
25	$A(S[2]) = X(S[2]) - X(S[3]);$
26	$L(S[2]) = \text{find_right}(X, S[2], S[1]) - S[3] + 1;$
27	$S = \text{pop}(S, [3]);$
28	end if
29	end if
30	$i := i + 1$
31	end for
32	// (4) Prop.4: Reversed terminal leg reduction
33	while $\text{len}S \geq 3$
34	$A(S[2]) = X(S[2]) - X(S[1]);$
35	$L(S[2]) = S[1] - \text{find_left}(X, S[2], S[3]) + 1;$
36	$S = \text{pop}(S, [1]);$
37	end while
38	return $[A, L]$

Figure 8: Maximal convex amplitude calculation

Line 1-6 initializes the output variables A and L and global variables S , P and M that are referred by a function "getNextLocalMaximum" described in Fig 9. . S is a LIFO (last in first out) stack. P and M are integers that store the time points.

```

Algorithm: getNextLocalMaximum (X,i)
Input: X: time series, i: time point
Global Var.: S, P, M
Output: next maximal time
01 if P == undef
02   return i;
03 else // P ≠ undef
04   while X[i] == X[P]
05     i := i + 1;
06   end while // X[i] ≠ X[P]
07   if S == undef
08     M := P;
09     return i;
10   else // X[i] ≠ X[P], M ≠ undef
11     while (X[i] - X[P])(X[i] - X[M]) > 0
12       P := i;
13       i := i+1;
14     end while // i is local maximal
15     M := P;
16     return i;
17   end if
18 end if
    
```

Figure 9: Subroutine getNextLocalMaximum

Line 7-31 is the main part that computes the maximal convex curves by scanning the time series from the beginning to the end.

Line 10-11 extracts a locally maximal time series from a given time series. That is, it transforms the given time series to the ones which consist of local maximum points (Fig 10).

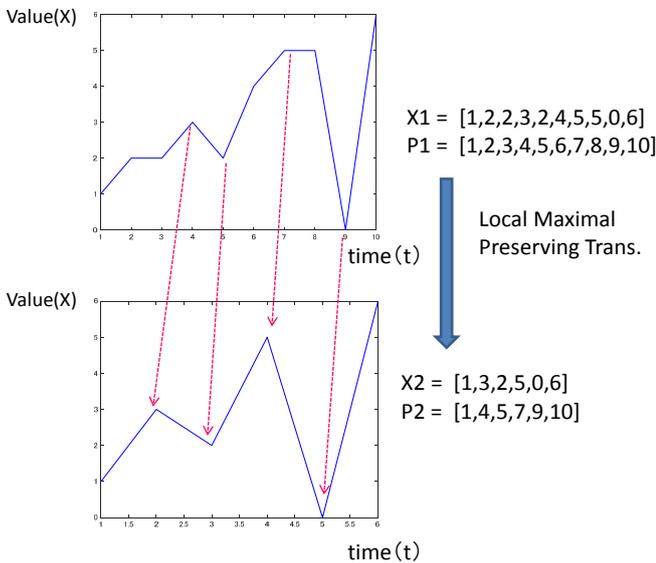


Figure 10: Local maximum preserving trans.

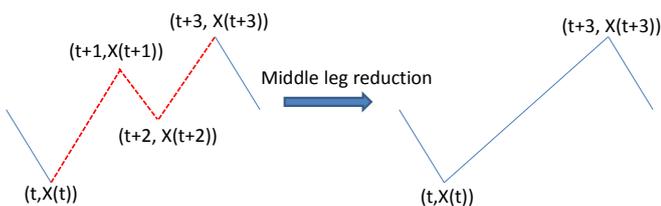


Figure 11: Middle leg reduction

Line 13-21 executes a middle leg reduction shown by Fig. 11.

A function "findleft" is a procedure that finds the left terminal of a maximal convex curve in Fig. 12. The worst case of computational complexity for finding the exact left terminal is O(n), where n is the length of time series. Therefore we accept an approximate solution which calculates the point of intersection of the line between the vertex (S[3],X(S[3])) and the left terminal (S[4], X(S[4])) and the line extended horizontally from the right terminal (S[2],X(S[2])). Its computational complexity is O(1). The length is not always necessary, so there is not a practical problem by calculating it when it is necessary.

The left terminal of balanced convex curve at (S[3], X(S[3]))

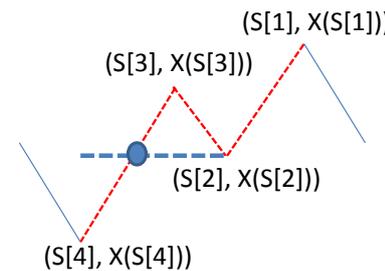


Figure 12: find_left(X, S[3], S[4])

Line 23-29 executes a terminal leg reduction shown by Fig. 13.

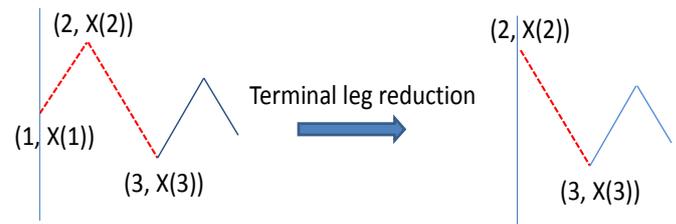


Figure 13: Terminal leg reduction

Line 33-37 executes a reversed terminal leg reduction shown by Fig. 14. It calculate the maximal amplitudes for the left time series, which is shown by Fig. 15., after middle leg reductions and terminal leg reductions are applied repeatedly until they cannot be applicable.

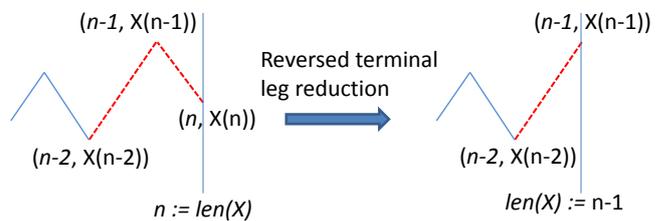


Figure 14: Reversed terminal leg reduction

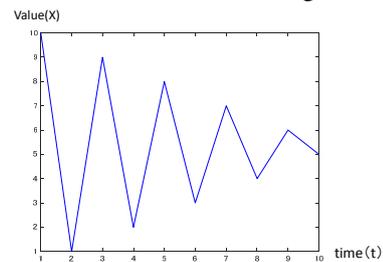


Figure 15: Shrinking time series

The computational complexity of the inside of the for-loop 08-31 and while-loop 34-37 are $O(1)$. The number of the for-loop is n and that of the while-loop is the size of the stack S that is at most n . Therefore, the total computational complexity of the proposed algorithm is $O(n)$, because the number of the reduction in while loops is at most n . The above algorithm is online, because the main loop is for-loop 08-31 with time and the while-loop 34-37 is a post-process to process the values that remain in the stack S after the for-loop 08-31.

4 EVALUATION

Section 2 shows the proposed new feature "the amplitude and the length of a maximal convex at each time" is parameter free. Section 3 shows the proposed algorithm to calculate the feature is online and its computational complexity is $O(n)$.

We implemented the proposed algorithm in section 3 with Matlab. This section shows that our proposed method can extract convex-shaped spikes in transient data by the experiment with simulated data.

(1) Noisy sine wave

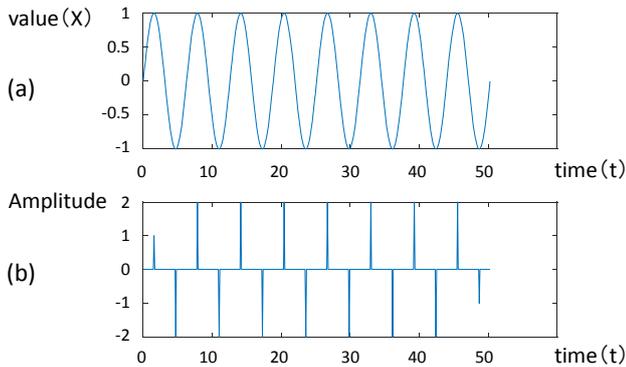


Figure 16: Convex amplitude of sine wave

Firstly, we explain the qualitative property of a maximal convex by showing the result for a noisy sine wave. In Fig. 16, (a) is a sine curve and (b) is a signed amplitude function that maps from a time point to the signed amplitude of the extended maximal convex at the time point. The amplitude at a local maximal point is the difference between the local maximal value and the next minimal value.

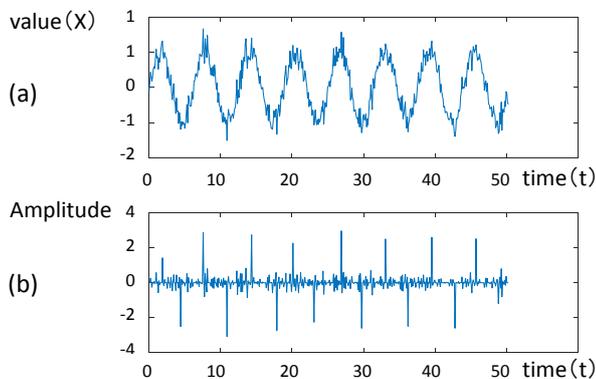


Figure 17: Convex amplitude of noisy sine wave

In Fig. 17, (a) is a sine curve with noise and (b) is a signed amplitude function for (a). The graph (b) shows that the amplitude at a local maximal point reflects the volume of a noise even during the transient period.

(2) Transient data with spike

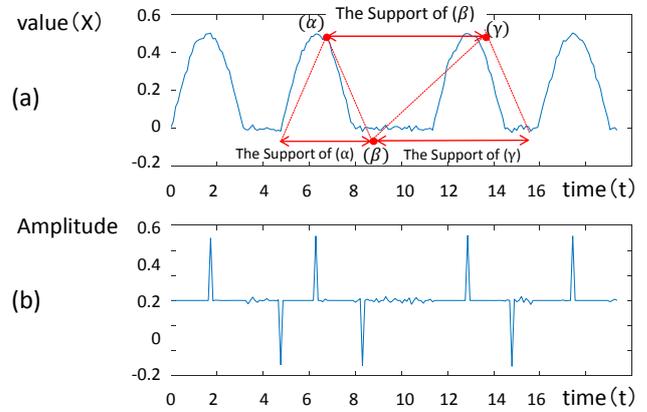


Figure 18: Transient data (normal)

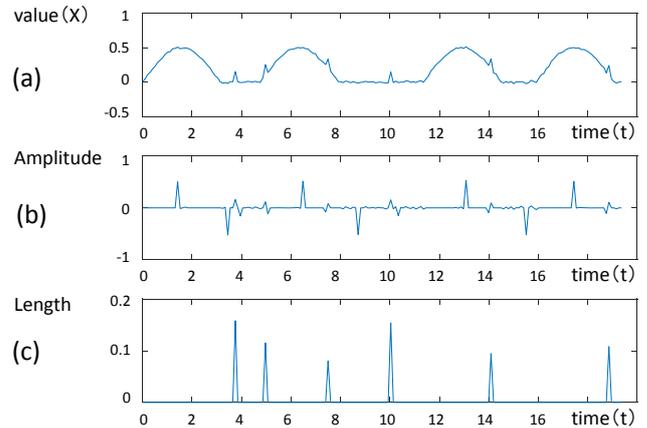


Figure 19: Transient data with spike (anomaly)

In Fig. 18, (a) is a normal transient trend and (b) is a signed amplitude function for (a). The graph (b) shows the scale of the amplitude of a normal operational pattern is 0.5. Note that there is a time that has a 0.5 scale amplitude during the steady period. If the supports of maximal convex curves intersect, we should select convexes the supports of which do not intersect. As in the case of Fig.17, we should select convex curves (α) and (γ) and discard (β) . We call (α) a real maximal convex and call (β) a virtual maximal convex curve.

In Fig. 19, (a) is an anomaly transient trend mixed with convex-shaped spikes and (b) is a signed amplitude function for (a). We can extract mixed convex-shaped spikes with 0.1 scale by deleting the scale of the amplitude that a normal trend has and deleting virtual maximal convex curves. The signed amplitude function (c) shows that the above filter can extract mixed-shaped spikes.

5 CONCLUSIONS

We proposed a new parameter-free fluctuation feature of time series that calculates the amplitude of a maximal

convex curve for extracting convex-shaped spikes in transient sensor data. We also shows an online algorithm whose computational complexity is $O(n)$, where n is the length of input time series, in order to be able to apply it to real time equipment condition monitoring.

Future works are the following.

(1) The proof of the proposed algorithm

We proposed a new algorithm and evaluated it empirically. To prove that our algorithm certainly calculates the maximal convex curve defined in section 2.1 is the first issue to be addressed in the future.

(2) Evaluation with real sensor data.

We showed the possibility of effectiveness of our algorithm in simulation data and the computational complexity of our proposed algorithm. The evaluation of its function and speed with real sensor data is an issue to be addressed in the future.

(3) Comparison of execution times with existing algoirthm

We analyzed the computational complexity of our algorithm. The comparison of execution times with that of Fink's algorithm, which is an existing algorithm, is an issue to be addressed in the future.

This work is supported by JSPS KAKENHI Grant Number 17K00161.

REFERENCES

- [1] J. Zheng, D. Simplot-Ryl, C. Bisdikian, H. T. Mouftah: "The Internet of Things [Guest Editorial]", Communications Magazine, IEEE, Vol.49, No.11, pp.30-31 (2011).
- [2] M. Imamura, D. Nikovski, Z. Sahinoglu, M. Jones: "A Survey on Machine Learning for Equipment Condition Monitoring Using Sensor Big Data", IIEEJ Transactions on Image Electronics and Visual Computing Vol.2 No.2, pp. 112-121 (2014).
- [3] E. Fink, B. P. Kevin: "Indexing of Compressed Time series", DATA MINING IN TIME SERIES DATABASES, World Scientific, pp. 43-65 (2004).
- [4] M. Imamura, T. Nakamura, H. Shibata, N. Hirai, S. Kitagami, T. Munaka: "Leg Vibration Analysis for Time Series", IPSJ Journal, vol. 57, No.4, pp.1303-1318 (2016). (in Japanese).
- [5] M. Imamura, D. Nikovski, M. Jones: An Anomaly Detection System for Equipment Condition Monitoring, International Journal of Informatics Society (IJIS) VOL.8, NO.3, pp. 161-169 (2016).

Session 8:
Industrial Applications
(Chair: Ryozo Kiyohara)

A Study on Optimum Mega-Photovoltaic System Design

Hironori Honda *, Yasushi Tachibana** , Jin Umeoka***,
Masashi Saito**** and Yasumitsu Suzuki*****

*Graduate School of Business Architecture, Kanazawa Institute of Technology, Japan

** Industrial Research Institute of Ishikawa, Japan

*** Shigemitsu Shoji Co., Ltd., Japan

**** Faculty of Kanazawa Institute of Technology, Japan

hironori@neptune.kanazawa-it.ac.jp

Abstract - We have analyzed factors required for efficient operation of mega-photovoltaic system to explore optimum system design from collected data. Collected data consist of generated power amount, solar radiation, outside air temperature, season and region since year 2012 when Feed-in Tariff (FIT) started in Japan.

First, we have analyzed the correlation among generated power amount, cumulative solar radiation and average outside air temperature by multiple regression. The results of multiple regression have shown that the influence of average outside air temperature to the generated power amount has been lower than the cumulative solar radiation.

Second, we have analyzed the impact on the generated power amount by the photovoltaic module angle from three different plants, north-facing 3 degree, south-facing 10 degree and south-facing 20 degree. We have confirmed that south-facing 20 degree has had both the most solar radiation and the most generated power amount.

As a result, we have convinced that the securing the generated power amount have been strongly influenced by cumulative solar radiation not by average outside air temperature. South-facing installation and angle are some factors for efficient operation of mega-photovoltaic system.

Keywords: photovoltaic power generation, generated energy, Amount of solar radiation, average outside air temperature, statistical analysis

1 INTRODUCTION

Because of the Great East Japan Earthquake, nuclear power plants in Japan have stopped for a long time.

As nuclear power plants occupied about 30% of the electricity used so far, the thermal power plants compensated for the sudden power shortage. As a result, the operation rate of thermal power plants with large fuel consumption increased, which led to the rise in electricity rates ^[1].

Because Japan relies on imports for most of its energy resources, it is affected not only by foreign situations, but also by how energy is used.

After the earthquake, issues such as safety of nuclear power plant, global warming of thermal power plant and generated energy cost have been pointed out, and renewable energy have become popular.

In order to promote the dissemination of renewable energy in July 2012, the Ministry of Economy, Trade and Industry, Japan implemented Feed-in Tariff (FIT). FIT focused on renewable energy generation businesses so that renewable energy has widely spread ^[1]. Variations in generated energy due to seasons and weather are one of the problems for photovoltaic power generation, especially for mega-photovoltaic system.

Conventional generated energy is indispensable in order to supply stable electric power at the electric power demand fluctuating throughout the day, and effective utilization of photovoltaic power generation is important to supply electricity making use of other power sources ^[5].

In this research, we report the relation between solar radiation amount and generated energy for the optimum design of mega solar based on actual generated energy data of 7 photovoltaic power plants installed in Ishikawa Prefecture, Japan.

2 SPECIFICATION AND LOCATION OF THE POWER PLANTS

This research is based on the data of photovoltaic power plants of Shigemitsu Shoji Co., Ltd., which is headquartered in Kanazawa city, Ishikawa prefecture, Japan. Shigemitsu Shoji Co., Ltd. has operated eight photovoltaic power plants in different locations. We show details in Table 1. In addition, the solar radiation meter of each generated energy is installed in accordance with the mounting surface of the photovoltaic module, and it is not installed horizontally with respect to the ground. Figure 1 show the connection status of PV1 and PV3 for reference.

Because the generated energy data used for the analysis has been up to May 31, 2016, the target is from PV1 to PV7.

Table 1: Installation details of photovoltaic power plant.

PV1 (Osaki, Kahoku-shi) 	generation scale		997.5kw/ 431.8kw	PV2 (Shichimi, Noto-cho) 	generation scale		752.5kw
	module	# of module	3,990/ 1,222		module	# of module	3,010
		direction	North/-			direction	South
		installation angle	3° /0°			installation angle	20°
	PSC		750kw/ 350kw		PSC		750kw
Consolidation date		2012.9.14/ 2015.8.6	Consolidation date		2013.3.25		
PV3 (Shimbomachi, Hakui-shi) 	generation scale		1,998.5kw	PV4 (Fujinami, Noto-cho) 	generation scale		994.0kw
	module	# of module	7,994		module	# of module	3,976
		direction	South			direction	South
		installation angle	20°			installation angle	20°
	PSC		2,000kw		PSC		1,000kw
Consolidation date		2013.3.22	Consolidation date		2013.9.10		
PV5 (Satohongo, Shika-machi) 	generation scale		979.3kw	PV6 (Futatsuya, Kahoku-shi) 	generation scale		1,271.6kw
	module	# of module	3,752		module	# of module	4,872
		direction	South			direction	South
		installation angle	20°			installation angle	10°
	PSC		750kw		PSC		1,200kw
Consolidation date		2012.9.14	Consolidation date		2014.8.29		
PV7 (Monzemmachi, Wajima-shi) 	generation scale		796.6kw	PV8 (Takidanimachi, Hakui-shi) 	generation scale		1,046.2kw
	module	# of module	3,052		module	# of module	3,948
		direction	South			direction	South
		installation angle	10°			installation angle	10°
	PSC		750kw		PSC		1,000kw
Consolidation date		2015.2.18	Consolidation date		2016.7.28		



Figure 1: Installation details of the solar radiation meter.

3 GENERATED ENERGY DATA ANALYSIS

Using data from PV1 to PV7, we have performed multiple regression analysis on how daily total solar radiation and

daily average outside air temperature are related to the photovoltaic power generation of the day.

Table 2 shows the analysis results of the first power plant (PV1). The adjusted R^2 of the regression statistics is 0.9774, indicating a high reliability result. According to this data, the coefficient of "PV1 integrated solar radiation amount" is 803.057, and the coefficient of "PV1 average outside air temperature" is 6.523.

From these results, it was confirmed that the cumulative solar radiation has an influence of about 50 times the average outside air temperature with respect to the cumulative generated energy amount of PV1. Likewise, multiple regression analysis was conducted from the second power plant (PV2) to the seventh power station (PV7). Those results are shown in Tables 3 to Tables 7. Making check these analysis results, the adjusted R^2 value at the 4th power station (PV4) is 0.6276, so they might include some data errors.

As a result, we confirmed that the influence of the average outside air temperature on the generated energy amount is much smaller than the cumulative solar radiation amount.

Table 2: Results of multiple regression analysis of cumulative solar radiation amount and average outside air temperature with respect to generated power amount of the PV1.

regression statistics	
Multiple Correlation Coefficient	0.989
Multiple R-Squared	0.977
Adjusted R-Squared	0.977
Standard Error	306.836
Observed Quantity	858,000

Analysis of Variance table					
	Degrees of Freedom	Fluctuation	Variance	Observed Dispersion Ratio	Significance F
Regression	2	3494438024	1.75E+09	18558.157	0
Residual	855	80496798.17	94148.3		
Total	857	3574934822			

	Coefficient	Standard Error	t Value	P-Ratio	Lower Limit 95%	Upper Limit 95%	Lower Limit 95.0%	Upper Limit 95.0%
Intercept	-40.181	22.645	-1.774	0.076	-84.628	4.266	-84.628	4.266
PV1 Cumulative Solar Radiation Amount (kWh/m ²)	803.057	5.040	159.332	0.000	793.165	812.950	793.165	812.950
PV1 Average Outside Temperature (°C)	6.524	1.430	4.561	0.000	3.716	9.331	3.716	9.331

Table 3: Results of multiple regression analysis of cumulative solar radiation amount and average outside air temperature with respect to generated power amount of the PV2.

regression statistics	
Multiple Correlation Coefficient	0.960
Multiple R-Squared	0.921
Adjusted R-Squared	0.921
Standard Error	458.363
Observed Quantity	1157,000

Analysis of Variance table					
	Degrees of Freedom	Fluctuation	Variance	Observed Dispersion Ratio	Significance F
Regression	2	2821405418	1.41E+09	6714.556	0
Residual	1154	242451005.7	210096.2		
Total	1156	3063856424			

	Coefficient	Standard Error	t Value	P-Ratio	Lower Limit 95%	Upper Limit 95%	Lower Limit 95.0%	Upper Limit 95.0%
Intercept	-11.276	29.255	-0.385	0.700	-68.674	46.123	-68.674	46.123
PV2 Cumulative Solar Radiation Amount (kWh/m ²)	661.618	6.461	102.403	0.000	648.942	674.295	648.942	674.295
PV2 Average Outside Temperature (°C)	3.872	1.902	2.035	0.042	0.139	7.604	0.139	7.604

Table 4: Results of multiple regression analysis of cumulative solar radiation amount and average outside air temperature with respect to generated power amount of the PV3.

regression statistics	
Multiple Correlation Coefficient	0.975
Multiple R-Squared	0.950
Adjusted R-Squared	0.950
Standard Error	989.711
Observed Quantity	1157.000

Analysis of Variance table					
	Degrees of Freedom	Fluctuation	Variance	Observed Dispersion Ratio	Significance F
Regression	2	21659039772	1.08E+10	11055.859	0
Residual	1154	1130374978	979527.7		
Total	1156	22789414750			

	Coefficient	Standard Error	t Value	P-Ratio	Lower Limit 95%	Upper Limit 95%	Lower Limit 95.0%	Upper Limit 95.0%
Intercept	309.790	64.701	4.788	0.000	182.844	436.736	182.844	436.736
PV3 Cumulative Solar Radiation Amount (kWh/m ²)	1779.956	13.332	133.513	0.000	1753.799	1806.113	1753.799	1806.113
PV3 Average Outside Temperature (°C)	-3.419	3.980	-0.859	0.391	-11.228	4.390	-11.228	4.390

Table 5: Results of multiple regression analysis of cumulative solar radiation amount and average outside air temperature with respect to generated power amount of the PV4.

regression statistics	
Multiple Correlation Coefficient	0.793
Multiple R-Squared	0.628
Adjusted R-Squared	0.628
Standard Error	1421.502
Observed Quantity	991.000

Analysis of Variance table					
	Degrees of Freedom	Fluctuation	Variance	Observed Dispersion Ratio	Significance F
Regression	2	3376605020	1.69E+09	835.517	3.9473E-213
Residual	988	1996419581	2020668		
Total	990	5373024601			

	Coefficient	Standard Error	t Value	P-Ratio	Lower Limit 95%	Upper Limit 95%	Lower Limit 95.0%	Upper Limit 95.0%
Intercept	540.604	90.216	5.992	0.000	363.567	717.641	363.567	717.641
PV4 Cumulative Solar Radiation Amount (kWh/m ²)	780.529	21.787	35.826	0.000	737.775	823.282	737.775	823.282
PV4 Average Outside Temperature (°C)	-1.540	6.253	-0.246	0.806	-13.811	10.731	-13.811	10.731

Table 6: Results of multiple regression analysis of cumulative solar radiation amount and average outside air temperature with respect to generated power amount of the PV5.

regression statistics	
Multiple Correlation Coefficient	0.963
Multiple R-Squared	0.926
Adjusted R-Squared	0.926
Standard Error	564.300
Observed Quantity	685.000

Analysis of Variance table					
	Degrees of Freedom	Fluctuation	Variance	Observed Dispersion Ratio	Significance F
Regression	2	2734818867	1.37E+09	4294.164	0
Residual	682	2171722255	318434.3		
Total	684	2951991092			

	Coefficient	Standard Error	t Value	P-Ratio	Lower Limit 95%	Upper Limit 95%	Lower Limit 95.0%	Upper Limit 95.0%
Intercept	20.412	46.423	0.440	0.660	-70.736	111.561	-70.736	111.561
PV5 Cumulative Solar Radiation Amount (kWh/m ²)	868.925	10.762	80.743	0.000	847.795	890.055	847.795	890.055
PV5 Average Outside Temperature (°C)	4.831	3.064	1.577	0.115	-1.184	10.847	-1.184	10.847

Table 7: Results of multiple regression analysis of cumulative solar radiation amount and average outside air temperature with respect to generated power amount of the PV6.

regression statistics	
Multiple Correlation Coefficient	0.982
Multiple R-Squared	0.965
Adjusted R-Squared	0.965
Standard Error	500.999
Observed Quantity	639.000

Analysis of Variance table					
	Degrees of Freedom	Fluctuation	Variance	Observed Dispersion Ratio	Significance F
Regression	2	4393700124	2.2E+09	8752.385	0
Residual	636	159636112	251000.2		
Total	638	4563336236			

	Coefficient	Standard Error	t Value	P-Ratio	Lower Limit 95%	Upper Limit 95%	Lower Limit 95.0%	Upper Limit 95.0%
Intercept	87.034	41.947	2.075	0.038	4.662	169.406	4.662	169.406
PV6 Cumulative Solar Radiation Amount (kWh/m ²)	1141.410	9.882	115.505	0.000	1122.005	1160.815	1122.005	1160.815
PV6 Average Outside Temperature (°C)	-1.502	3.057	-0.491	0.623	-7.504	4.501	-7.504	4.501

Table 8: Results of multiple regression analysis of cumulative solar radiation amount and average outside air temperature with respect to generated power amount of the PV7.

regression statistics	
Multiple Correlation Coefficient	0.979
Multiple R-Squared	0.959
Adjusted R-Squared	0.958
Standard Error	339.286
Observed Quantity	457.000

Analysis of Variance table					
	Degrees of Freedom	Fluctuation	Variance	Observed Dispersion Ratio	Significance F
Regression	2	1208625904	6.04E+08	5249.639	0
Residual	454	52262272.21	115115.1		
Total	456	1260888176			

	Coefficient	Standard Error	t Value	P-Ratio	Lower Limit 95%	Upper Limit 95%	Lower Limit 95.0%	Upper Limit 95.0%
Intercept	27.900	37.030	0.753	0.452	-44.872	100.673	-44.872	100.673
PV7 Cumulative Solar Radiation Amount (kWh/m ²)	715.414	7.922	90.310	0.000	699.846	730.982	699.846	730.982
PV7 Average OutsideTemperature (°C)	7.471	2.464	3.032	0.003	2.629	12.313	2.629	12.313

4 CORRELATION ANALYSIS ON GENERATED ENERGY

From the results in Chapter 3, we have confirmed that the amount of solar radiation has a major impact on generated energy. However, the amount of solar radiation changes not only according to latitude, but also depending on azimuth angle and inclination angle [3]. As described in Chapter 2, the angles of the modules and installation locations differ depending on each power plant. The positional relation of the power plants is as shown in Figure 2.

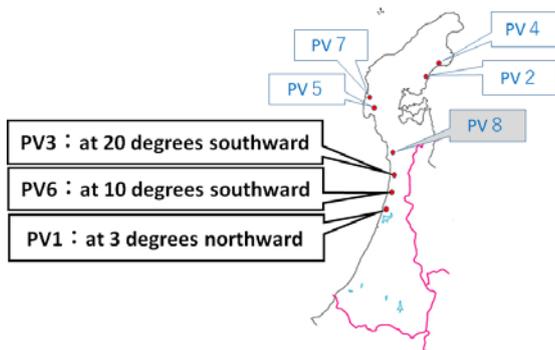


Figure 2: Position of each power plant.

Among them, performed Correlation analysis using solar radiation amount and power generation amount data with shorter distances between power stations and different installation angles, PV1, PV3 and PV6. The period of used data is following:

- PV1 is between 2013/4/1 and 2015/8/6,
- PV3 is between 2013/4/1 and 2016/5/31,
- PV6 is between 2014/9/1 and 2016/5/31.

The photovoltaic module angle installed on PV1 is almost flat with its inclination angle 3 degrees northward. From 2015/8/7, photovoltaic modules were added and the generated energy amount changed, so data up to that time has been used. The photovoltaic module of the PV3 is installed at 20 degrees southward and the photovoltaic module of the PV6 is installed at 10 degrees southward.

Also, since the consolidation date of PV6, which is installed lately, is August 29, 2014, the subsequent data has been used. We show correlation analysis results of generated energy amount - solar radiation amount and generated

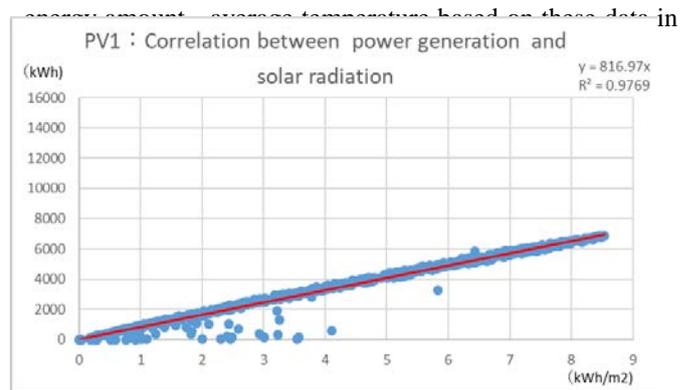


Figure 3: Correlation between generated energy and temperature (PV1).

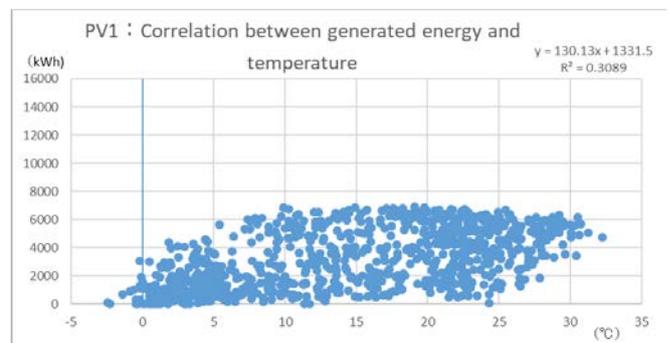


Figure 4: Correlation between generated energy and temperature (PV1).

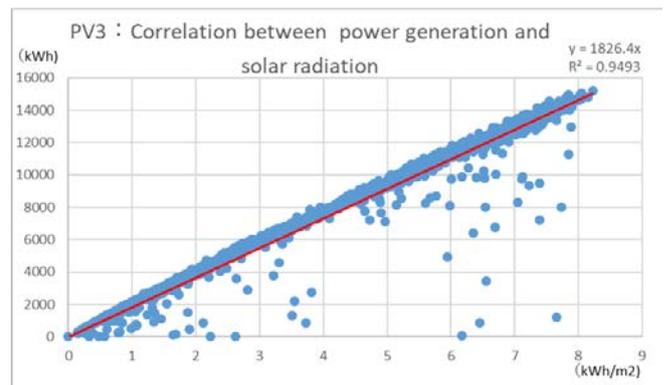


Figure 5: Correlation between generated energy and solar radiation (PV3).

PV6 : $y=1154.3x$
 PV1 : $y=817.0x$

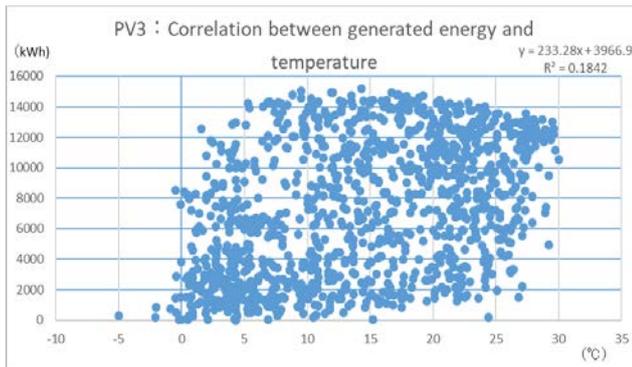


Figure 6: Correlation between generated energy and temperature (PV3).

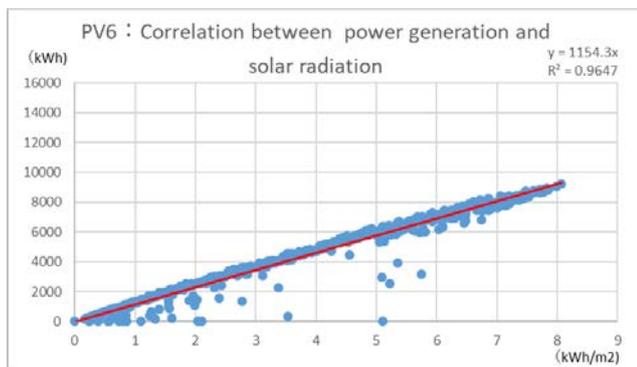


Figure 7: Correlation between generated energy and solar radiation (PV6).

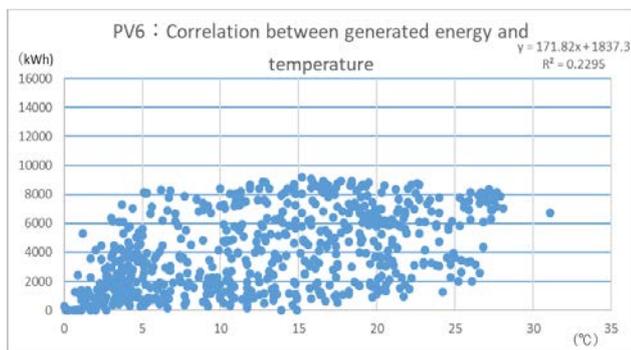


Figure 8: Correlation between generated energy and temperature (PV6).

In the correlation analysis between the generated energy amount and solar radiation, the correlation coefficient of PV1, PV3, and PV6 exceed 0.949. However, in the correlation analysis between the generated energy amount and the average temperature, the correlation coefficient of PV 1, PV 3, PV 6 is smaller than 0.31.

From these results, we are sure that there is almost no correlation between average temperature and generated energy.

When comparing the value of "solar radiation amount for generated energy amount", each PV has a high correlation result as follows:

PV3 : $y=1826.4x$

In this way, the difference can be seen in the value. However, even at the same solar radiation dose, the generated energy would be large at power plants with high generated energy capacity, so this result does not mean the difference in generated energy due to module angle. Therefore, it is necessary to compare the difference in generated energy amount by equalizing the generated energy capacities as follows:

PV1 : $816.97/997.5kw=0.819$
 PV3 : $1826.4/1998.5kw=0.914$
 PV6 : $1154.3/1271.6kw=0.908$

According to above values, the magnitude of the generated energy amount relation is as follows:

PV3 > PV6 > PV1

(South 20 degrees > South 10 degrees > North 3 degrees)

5 DIFFERENCE IN SOLAR RADIATION AMOUNT DUE TO DIFFERENT ANGLES

From the multiple regression analysis and the correlation analysis so far, we have shown that the generated energy amount is not substantially influenced by the average temperature, and strongly affects the amount of solar radiation. Therefore, from the solar radiation data of each power plant, we have investigated the relationship between the photovoltaic module angle and the solar radiation amount.

First, a comparison on a graph based on solar radiation data of PV 1 to PV 7 during the period from April 1, 2013 to May 31, 2016 is shown in Figure 9. From Figure 9, we can see that the amount of solar radiation increases or decreases repeatedly for every power station every season.

In order to compare the difference in solar radiation amount from the angle of the photovoltaic module, PV 1, PV 3 and PV 6 where the position of the power plant is close as shown in Figure 2 and the inclination angle is different are compared in the same way as in this chapter as shown in Figure 10 .

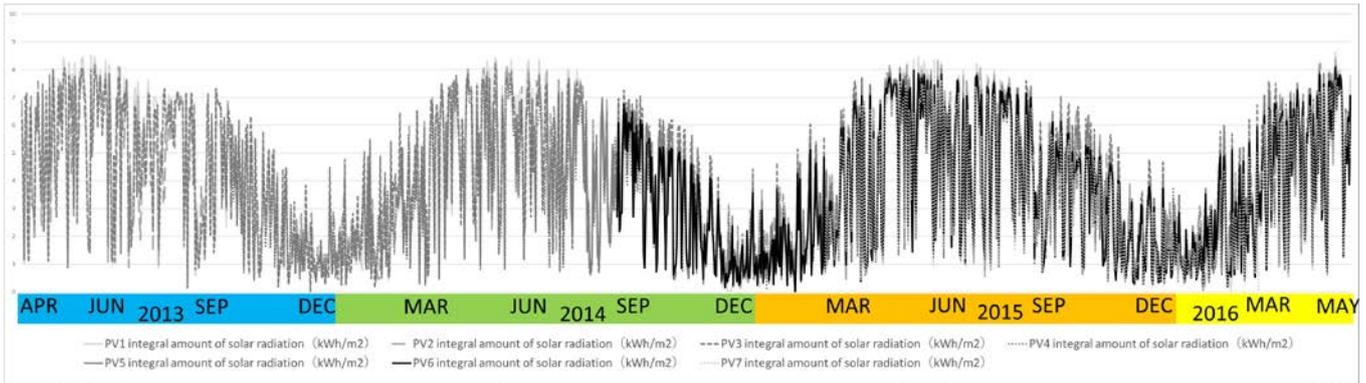


Figure 9: Comparison of solar radiation amount for PV1~ PV7. (Term: 2013.4.1~2016.5.31)

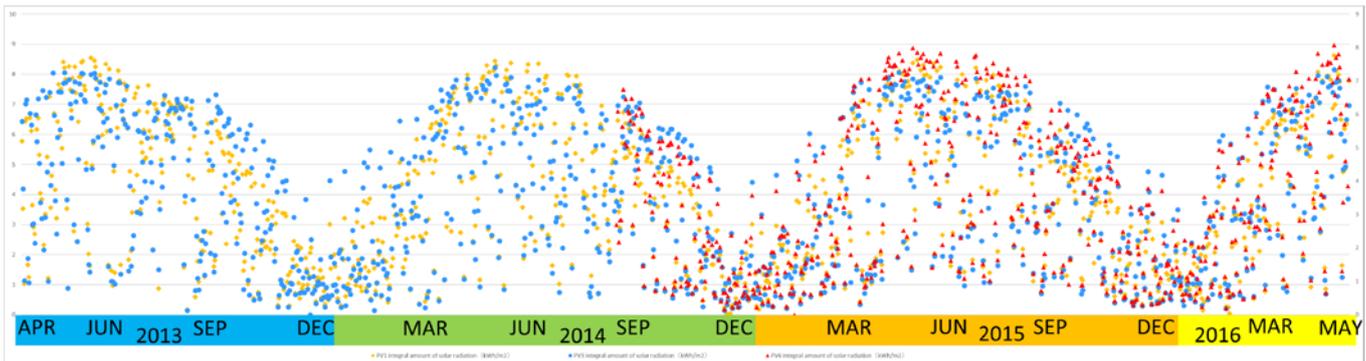


Figure 10: Comparison of solar radiation amount for PV1 (◆), PV3 (●), PV6 (▲). (Term: 2013.4.1~2016.5.31)

From Figure 10, we can identify that the square-mark PV1 (north 3 degrees) is shifting in a form slightly smaller than the other colors. We can also identify that PV3 (South 20 degrees) is larger than other colors. We can also identify that the difference of solar radiation is small in summer and the difference is widening in winter. So taking advantage of the characteristics of irradiation time, we divide the year into four and compare cumulative solar radiation amounts. We show the comparison results in Table 9.

Table 9: Comparison of accumulated solar radiation for PV1, PV3 and PV6.

	(kWh/m ²)		
	PV1	PV3	PV6
May~Jul	515.1916	490.289	484.435
Aug~Oct	382.8076	411.958	382.046
Nov~Jan	144.7966	171.867	152.563
Feb~Apr	336.3573	362.822	336.13
Total	1379.1531	1436.936	1355.174

As shown in Table 9, in May - July, when the sunshine hours are the longest throughout the year, PV1 with almost flat photovoltaic module angle gained the largest amount of solar radiation. In addition, we have found out that PV3 with a large installation angle of photovoltaic modules has the largest cumulative solar radiation from August to April, and the cumulative solar light volume per year is also the largest.

6 ABNORMAL VALUES OF GENERATED ENERGY DATA

When examining relationship analysis data of solar radiation amount and generated energy amount at each power station, it includes abnormal values. Especially from the results of multiple regression analysis and correlation analysis, R² of PV4 is low and the value showing reliability is low.

In the multiple regression analysis results in Table 3, adjusted R² is 0.627, and the correlation analysis result of PV4 generated power and solar radiation in Figure 11 also shows that R² is 0.6123 and the reliability of the data is low.

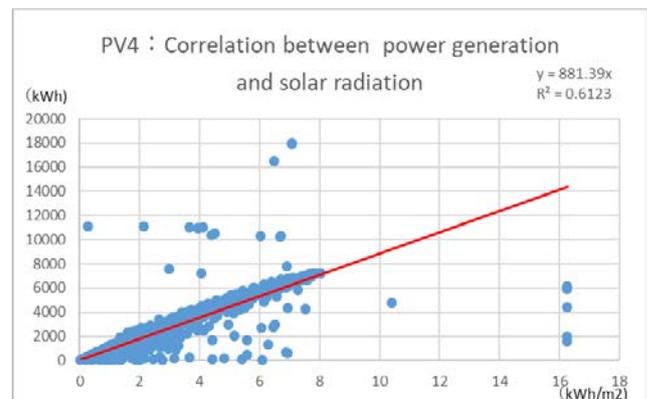


Figure 11: Correlation between generated energy and solar radiation (PV4).

Likewise, PV2 near PV4 has also low R^2 as shown in Figure 12, and the result of multiple regression analysis in Table 3 also has a low adjusted R^2 of 0.921. Additionally, for PV5, the R^2 shown in Figure 13 and adjusted R^2 of multiple regression analysis in Table 10 are as low as 0.926.

Table 10: Multiple regression analysis results for PV1 to PV7

	Adjusted R-Squared	Cumulative Solar Radiation Amount (kWh/m ²)	Average Outside Temperature (°C)
PV1	0.977	803.057	6.524
PV2	0.921	661.618	3.872
PV3	0.950	1779.956	-3.419
PV4	0.628	780.529	-1.540
PV5	0.926	868.925	4.831
PV6	0.965	1141.410	-1.502
PV7	0.958	715.414	7.471

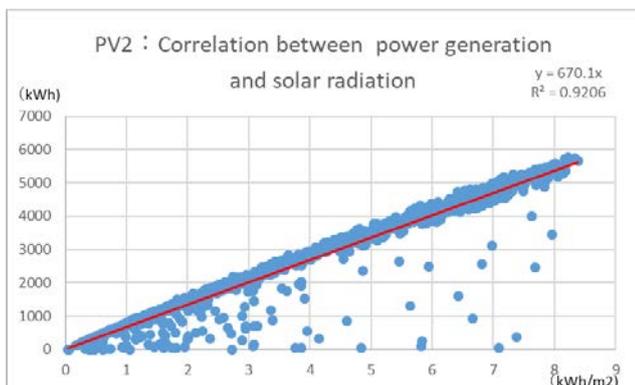


Figure 12: Correlation between generated energy and solar radiation (PV2).

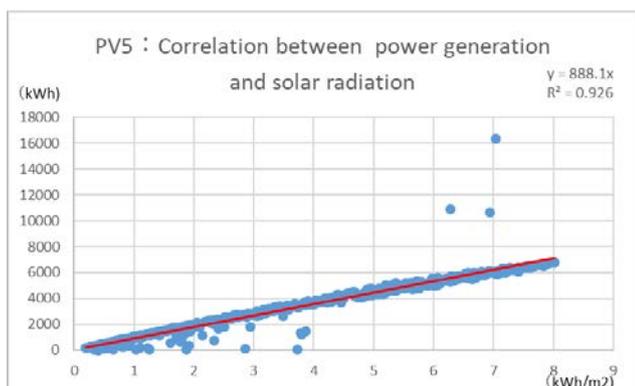


Figure 13: Correlation between generated energy and solar radiation (PV5).

We have found some abnormal generated energy values, i.e. even if the amount of solar radiation exists, there is no or little generated energy. On the contrary, some of them indicate that a large amount of electricity is generated despite the insolation level being small. We are assuming these abnormal values would be attributed to the influence of snow in winter and the problem of system voltage are

raise. For example, if snow is attached to a solar cell module, there is a possibility that electricity is not generated despite the amount of solar radiation [2].

Also, due to the problem of commercial power supply voltage, it may not be possible to generate electricity by safety control. This is due to the control that the power conditioner lowers the voltage of the photovoltaic power generation side to protect the power generation system when the electric power demand in the area to be power-supplied does not exist or the electric power supply at the time when at the electric power demand was low. These situation may occur at times. This is applicable even when electricity is supplied in areas with low population density and low electricity being used.

Table 11 shows the result of examining the population and number of households near the power plant.

Table 11: Population around the power station and number of households.

generating-station	address	population	households	Remarks
PV1,PV6	Kahoku-shi	35,060	12,632	2017/5/1
PV3	Hakui-shi	22,238	8,536	2017/5/1
PV7	Wajima-shi	28,051	12,690	2017/3/1
PV5	Shika-machi	21,086	8,074	2017/4/30
PV2,PV4	Noto-cho	18,076	7,753	2017/5/1

Comparing the multiple regression analysis results and correlation analysis results of each power plant with the surrounding population and households, we have found that PV1, PV6, PV7, which had higher adjusted R^2 values, had more population and households than other areas.

Meanwhile, it can be inferred that Noto town where PV2 and PV4 are installed have relatively few population and households, and these are areas where electricity demands are low. From this, we are assuming that abnormal values of PV2 and PV4 have problems of system voltage in addition to the effect of snow in winter. This means that when installing a photovoltaic power generation system, it is necessary to grasp the power demand situation around the installation site.

7 CONCLUDING REMARKS

In this paper, we have aimed to investigate optimum design that can effectively generate electricity as actual operating data by using mega photovoltaic power generation data accumulated for a long time. It is effective to set the installation angle of the photovoltaic module to the south side which is generally set. In the case of north / 3 degrees, south / 10 degrees, south / 20 degrees, it was confirmed that installation at south / 20 degrees has large amount of solar radiation and the amount of electricity generation more.

Moreover, for power generation, we are sure that the influence of the average temperature of the day hardly exists, and the cumulative solar radiation amount is important.

However, at the present stage, not all data have been analyzed, such as generated energy analysis excluding abnormal values and generated energy data per minute.

In the future, we will investigate these data as well as compare optimal inclination angle at power plant site with optimum inclination angle released by NEDO and to verify. Also, we plan to proceed with optimal design in areas that can be installed considering the influence of shadows.

REFERENCES

- [1] Ministry of Economy, Trade and Industry. Agency for Natural Resources and Energy, Annual report on energy 2016.
- [2] K. Hosokawa, and T. Tomabechei, Reseach on the Installation Technology of Photovoltaic Systems in Consideration of Snow, AIJ Journal of Tech. and Design, Vol.16, No.32, pp. 197-200 (2010).
- [3] K. Soichiro, S. Kazuhiro, and A. Hiroshi, A Study on Generated energy Characteristics of Photovoltaic Cells by Orientation and inclination angle, AIJ Kyushu Chapter Research Meeting, Vol. 40, pp. 353-356 (2001).
- [4] Y. Eiichi, Y. Toshiaki, T. Syuichi, A. Hiroshi, S. Kazuhiro, Estimation of output by different azimuth and inclination angle of photovoltaic cell, Proceedings of the Japan Society of Mechanical Engineers, Lecture Meeting of the 55th General Meeting of Kyushu Branch, No.028-1, 2002-03-18.
- [5] Environment Business Post FIT Special issue, Japan Business Publishing, pp. 16-20 (2017).
- [6] Shigemitsu Shoji Co., Ltd., Photovoltaic System, <http://www.shigemitsu-shoji.co.jp/PV/TOP.html>. (2017-04-15).
- [7] Ministry of Economy, Trade and Industry. Agency for Natural Resources and Energy, Feed-in Tariff (FIT) Guidebook2017, (2017-03).
- [8] NEDO Solar radiation database browsing system, app0.infoc.nedo.go.jp/, (reference: 2017-04-15).
- [9] Japan Atomic Technology Association. Nuclear power plant facility operation status, <http://www.gengikyo.jp/facility/powerplant.html>, (reference: 2017-04-13).
- [10] Photovoltaic power generation comprehensive information / Eco Life .com, Installation angle of Photovoltaic power generation and power generation efficiency, standard-project.net/solar/angle.html, (reference: 2017-04-13).
- [11] Municipal office of Noto-cho HP, http://www.town.noto.lg.jp/www/normal_top.jsp/, (reference: 2017-05-25).
- [12] Municipal office of Shika-machi HP, <http://www.town.shika.lg.jp/>, (reference: 2017-05-25).
- [13] Municipal office of Wajima-shi HP, <http://www.city.wajima.ishikawa.jp/docs/2017050900011/>, (reference: 2017-05-25).
- [14] Municipal office of Hakui-shi HP, <http://www.city.hakui.ishikawa.jp/sypher/www/info/detail.jsp?id=680>, (reference: 2017-05-25).
- [15] Municipal office of Kahoku-shi HP, <http://www.city.kahoku.ishikawa.jp/www/>, (reference: 2017-05-25).

A Trial of Ambient Sensor Method for Worker States Classifier

Kazuyuki Iso[†], Minoru Kobayashi[‡], and Takaya Yuizono[†]

[†]Japan Advanced Institute of Science and Technology, 1-1 Asahidai, Nomi, Ishikawa, Japan

[‡]Meiji University, 4-21-1 Nakano, Nakano-ku, Tokyo, Japan

{iso.kazuyuki, yuizono}@jaist.ac.jp
minoru@acm.org

Abstract - Sharing worker states using physical assets such as paper and whiteboards is assumed to be favorable for collaboration among workers in remote locations. Various methods to classify worker states have been studied in the past. The methods can be classified into three categories: ambient sensor method, wearable sensor method, and PC log method. In this study, we investigate the ambient sensor method using sensors that are compact and easy to integrate within the environment. Using these sensors, we can detect various kinds of information about the worker. A prototype of the ambient sensor method that uses a combination of a vibration sensor and a distance sensor was developed and tested. Self-organizing map (SOM), a machine learning technique was applied to the data obtained from both the sensors to classify four physical states of the remote worker: writing on a desk, typing on the PC with a keyboard, viewing a PC monitor, and leaving a seat. The study showed that the prototype module classified the worker states into one of the four states with an accuracy of 66%.

Keywords: Remote collaboration, Worker's state, Vibration sensor, Distance sensor, Self-organizing maps.

1 Introduction

Broadband networks enable Internet connectivity that allows multiple devices to connect and share multimedia content. This helps build a social infrastructure with which people can collaborate using multipoint video conferencing systems, where they can easily use real-time video communication and also share documents. For remote collaboration, recognition of a remote worker's state is an important factor. Research related to this factor has traditionally focused on awareness support [1], [2], [3], [4].

This research proposes a remote collaboration system that can utilize various physical assets such as paper and whiteboard. Workers who are co-located often use physical assets to collaborate. Therefore, it is assumed that using these physical assets and the sharing of physical user states are factors enabling effective collaboration between remote workers.

Various methods have been investigated to estimate worker states, but each method has some drawbacks. In this research, a vibration sensor and a distance sensor are proposed, prototyped, and tested. The vibration sensor detects signals based on worker behavior. The distance sensor detects the area in which the worker is present. Each sensor is extremely compact, easy to integrate with the environment, and can detect various kinds of information about the worker. With the emer-

gence of Internet of Things (IoT), the ambient environment with a combination of sensors can be used to develop an advanced collaboration infrastructure.

2 Related Research

2.1 Sensor approach

In this research, a method using a combination of vibration sensor data and distance sensor data to classify worker states is investigated by prototyping and testing. This method cannot be a standalone alternative to the proposed method but may supplement other methods in the IoT environment.

The ambient sensor method:

Sensors embedded in the work environment have been utilized in past studies to estimate worker states. Kennedy et al. [5] installed a microphone in the environment to estimate the worker states during a conference. However, the microphone was not very effective as it could not assess the involvement of the participants who were not taking part in the discussion.

Otsuka [6] reconstructed the state of the speaker and that of the listener using a video image. The camera was able to obtain various kinds of information but it was difficult to adjust this method to various conditions. In order to classify the state of the worker, the image of the worker must be of an appropriate size and angle. Moreover, the method did not work in instances of physical occlusion.

The wearable sensor method:

Murao et al. [7] proposed a method to assess remote worker states by attaching various sensors to the workers' bodies. The workers were made to wear acceleration sensors, which could capture detailed information about the remote worker and help evaluate their effort. However, attaching the sensor to users could be a major hindrance in remote collaboration.

The PC log method:

A personal computer's (PC) operational log is utilized to estimate a user state. Hashimoto et al. [8] obtained useful information about remote workers by maintaining an operational log of the workers' personal computers (PCs). However, the method did not differentiate between each users state and the log in case the workspace was shared.

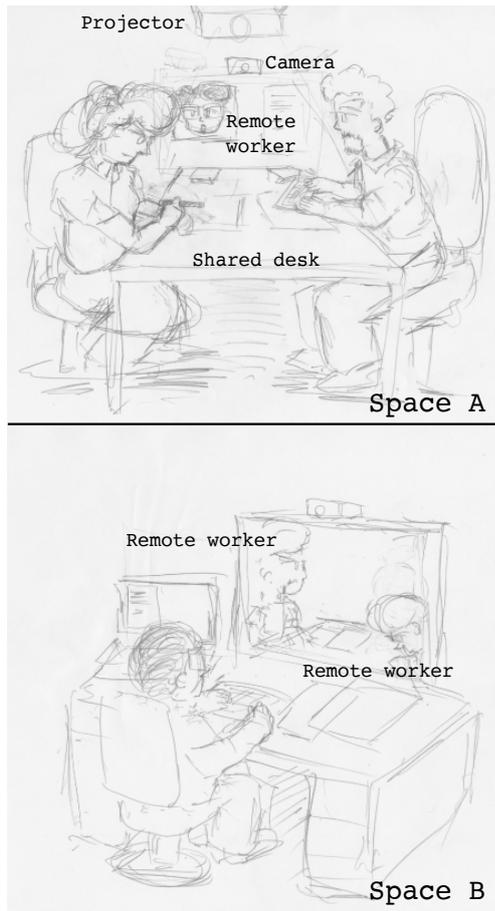


Figure 1: Concept of Remote Collaboration.

2.2 Realistic 3D Display approach

To support remote collaboration, a method was proposed that captures face and direction of the remote worker's gaze using a 3D display [9], [10], [11]. This method was expected to effectively support the level of awareness of remote workers. However, this method also has two drawbacks: first, it requires a very large capacity network for video transmission, and second, it requires many input devices, which are not easy to set up for generation of videos. Therefore, it is advisable to adopt an approach that does not involve the use of video to assess the awareness level of remote workers.

3 Remote Collaboration and Worker State Classifier

3.1 Remote Collaboration

To enhance remote collaboration when workers are located in different physical locations as depicted in Figure 1, we need a better method for sharing physical states.

In the illustration of remote collaboration shown in Figure 1, a group of people undertakes a task in a conference room in space A, while a worker in space B participates in the task remotely. In this situation, workers in space A may focus only on the discussion among them and forget or ignore the presence of the worker in space B, thereby excluding the remote

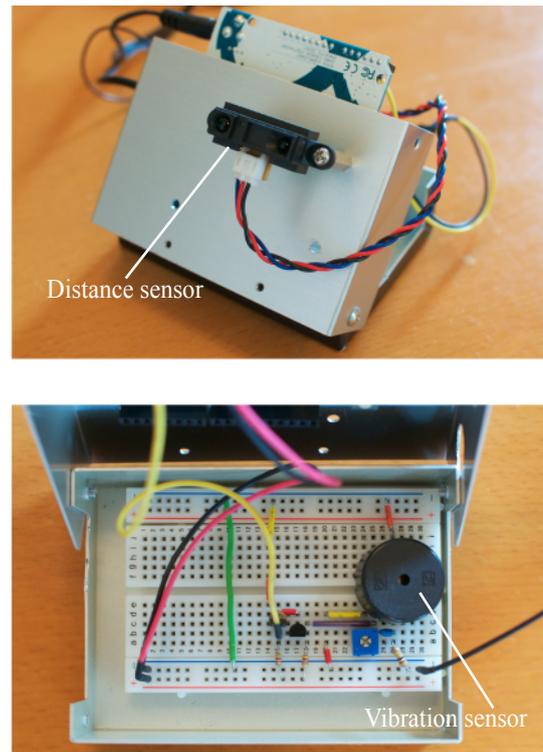


Figure 2: Prototype module with a combination of vibration and distance sensors.

worker from the discussion. To resolve this problem, a remote collaboration system needs to ensure that the workers in space A automatically share their working states with the remote worker in space B. If the workers in space A are continuously made aware of the presence of the remote worker, the collaboration will proceed more smoothly.

3.2 Worker State Classifier

In this study, we prototyped a worker state classifier that uses ambient sensors to assess worker state by combining the effect of vibrations and distance sensors. The ambient module is used by simply placing it on a desk (Figure 2). This module includes one vibration sensor and one distance sensor and receives data from both sensors synchronously using an Arduino microcontroller. The data is then transmitted to and stored on a PC.

We use self-organizing map (SOM) to classify worker states by analyzing the sensor data. SOM is a classification method that does not require teacher data. If teacher data is not required, a classification model corresponding to the work environment can be easily created.

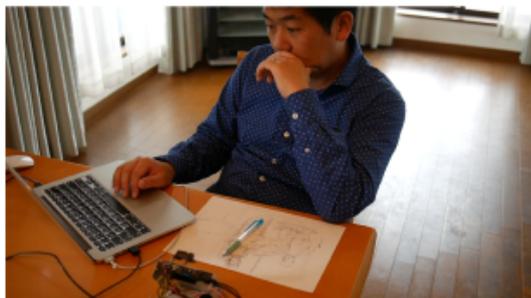
A Fourier transform is applied to the vibration sensor data and an average of the distance sensor data over a short period of time is taken into account for the data analysis. The SOM analyzes the relationship between the data from the sensors and the state of the worker, and the relationship can be applied for the classification of worker states.



Writing on a desk



Typing using the keyboard



Viewing a PC monitor

Figure 3: Worker State tested using the Prototype Module.

4 Worker State Classifier Procedure

4.1 Test Procedure

The prototype module was tested for its ability to deduce four states of the remote worker: state 1 of writing on a desk, state 2 of typing using the keyboard, state 3 of viewing a PC monitor, and state 4 of leaving a seat. The first three states are depicted in Figure 3. These four states were selected from the typical operations involving the use of a PC and physical assets.

During the 20-min test, a user worked at a desk and left the desk twice. The sampling rate of the two sensors was 1 kHz. The user was video recorded in order to label the sensor data

State 1: Writing on a desk

The worker makes a figure, scribbling some words and drawing lines. While drawing, the worker stops drawing and takes some time to think for a few seconds.

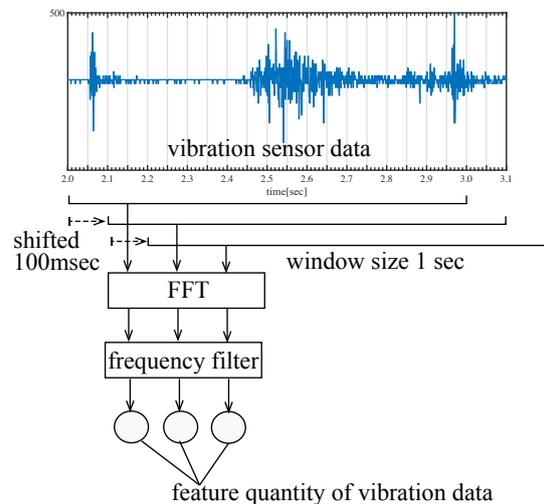


Figure 4: Method of calculating feature quantity using data obtained from the vibration sensor.

In case remote collaboration, figures made of letters, words, and simple lines are often drawn on paper or whiteboards. The properties of the vibrations recorded by the sensor are similar to those of co-located workers.

State 2: Typing using the keyboard

The worker creates sentences using a text editor on the PC. In the case of remote collaboration, a worker types a key and creates a report or minutes of meetings. The vibration of work occurring in such a scenario is similar to the work done by co-located workers.

State 3: Viewing a PC monitor

The worker browses the web by operating the touch pad. During this time, almost no vibration is observed by the sensor.

State 4: Leaving a seat.

The worker leaves the seat and moves to the next room. After a few minutes, the worker returns and sits on the seat. A significant vibration is recorded when the worker moves the chair.

The method of calculating the feature quantity is illustrated in Figure 4. The sensor data was analyzed using Fast Fourier Transform (FFT). The time window for recording the FFT was 1s. The feature quantity was calculated while shifting the window of FFT by 100ms. Finally, the feature quantity was passed through a frequency filter. The frequency filter was set to pass 2-200Hz.

The feature quantity of the distance sensor was calculated using the moving average of the sensor data. The moving average of the distance sensor was found to be 1s. The input of SOM was a combination of the two feature quantities.

4.2 Results

The graphs plotted from the data obtained by both the sensors are shown in Figure 5(a) and Figure 5(c). The vibration data labeled with the worker states is shown in Figure 5 (a).

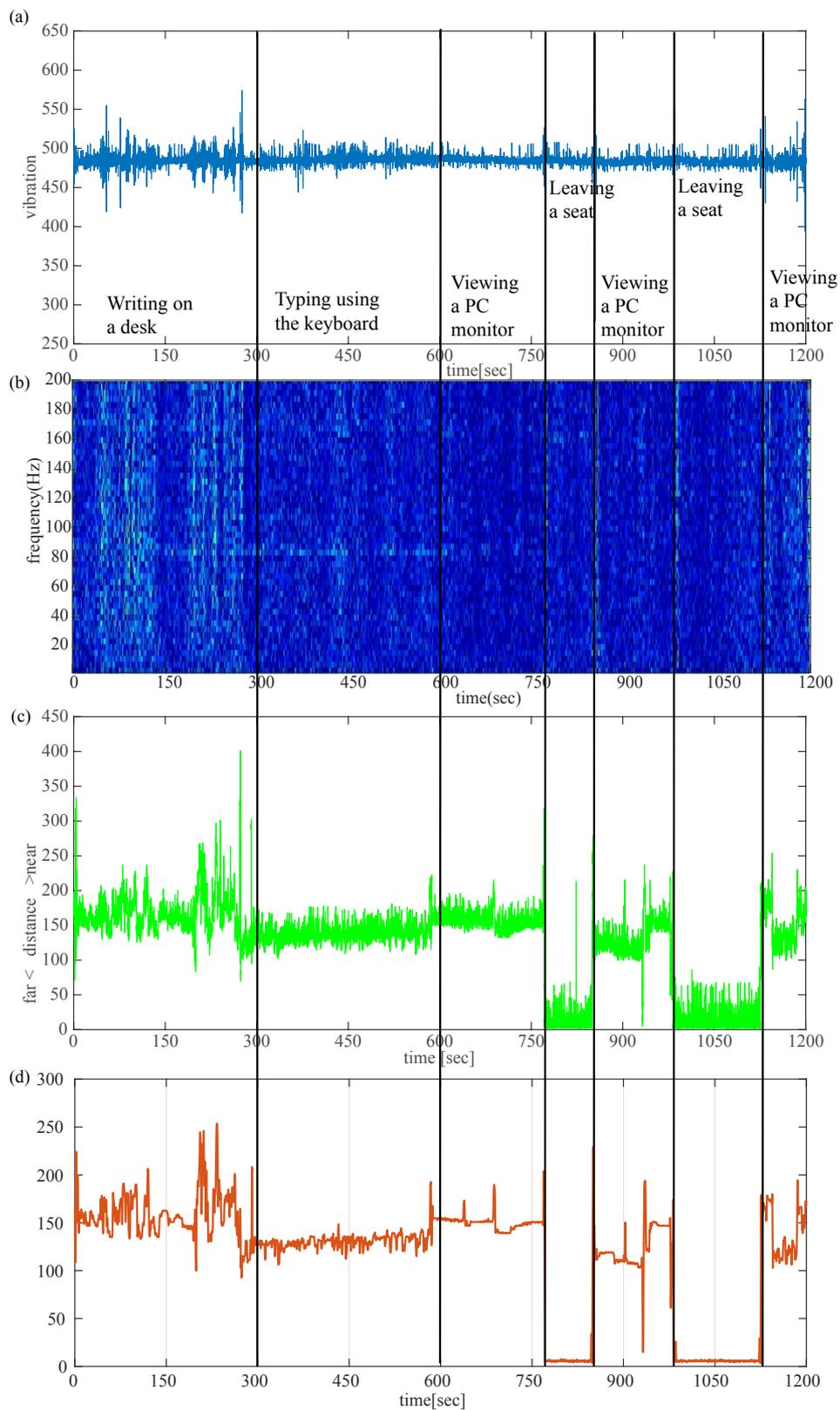


Figure 5: (a) Data obtained from vibration sensor depicting worker states (The Y-axis is a 10-bit value of an Arduino with an analog-to-digital converter) (b) Results of a short-time Fourier transform of the vibration sensor (c) Data obtained from distance sensor (d) Moving average of the distance sensor.

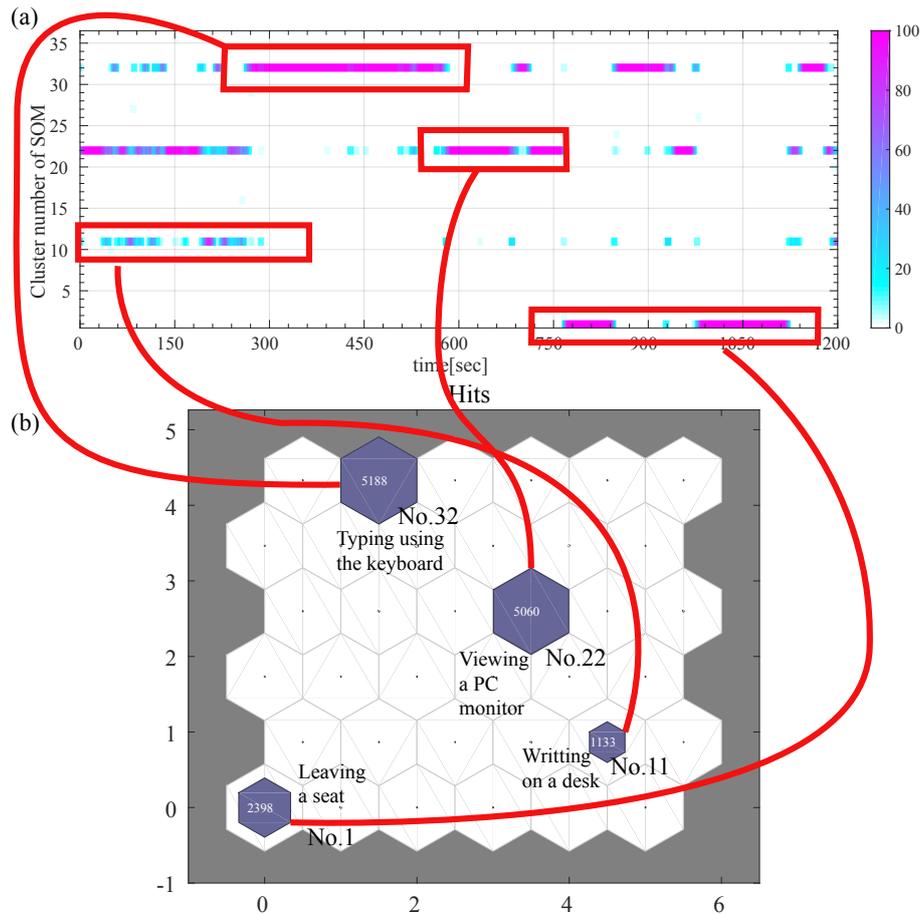


Figure 6: (a) Cluster number histogram of SOM calculated at 1 s intervals (b) SOM results obtained from the analysis.

The results of the short-time Fourier transform of the vibration sensor are shown in Figure 5 (b). The results of the moving average of the distance sensor are shown in Figure 5 (d).

The cluster number histogram of the SOM calculated at 10s intervals is shown in Figure 6 (a). The result of the SOM is shown in Figure 6 (b). The SOM classified the sensor data into four clusters. The rate of four worker status included in each cluster was found to be 66%.

The three states ‘typing using the keyboard,’ ‘viewing a PC monitor,’ and ‘leaving a seat’ are classified into cluster numbers different from each other as No. 1, No. 22, and No. 32. On the other hand, ‘writing on a desk’ is considered a part of two clusters and hence classified into two cluster numbers, No. 11 and No. 22. The time which was classified only in No. 22 was the state of ‘viewing a PC monitor’. The state that was classified in both No. 11 and No. 22 was the state of ‘writing on a desk’. Based on these results, the proposed method can identify four worker states.

5 Discussion

On comparing Figure 5 (b) with Figures 6(a) and (b), it becomes evident that the vibration data was related to the state of ‘typing using the keyboard.’ Therefore, it can be inferred

that the difference in vibration properties impacted the classification of No. 11, No. 22, and No. 32. In order to classify these three states, the feature quantity of the vibration sensor was adjusted using two parameters. The first is the time window of the FFT, and the second is the range of the frequency filter. While these two parameters were experimentally determined in this study, in the future, optimal values of these parameters can be generated by investigating the frequency characteristics of the transmitted vibration.

Figure 5 and Figure 6 show that the distance data was related to the state ‘leaving a seat.’ Therefore, the distance sensor influenced the classification, leading it to No. 1. These results suggest that the two sensors could be useful in classifying worker states.

In the case of remote collaboration, the workers are located in different physical spaces. These physical spaces are composed of various objects and these objects are made of different materials. The data obtained by the vibration sensor varies depending on the material that transmits the vibration. The value recorded by the distance sensor is related to the position of the object in the physical space. The relationship between the sensor data and the state of the worker, taking into account the state of the working environment, can also be investigated.

The three past studies described in the ‘Related Research’ section of this study were conducted under different environmental conditions. Depending on the components of the co-located workspace, the combination of these approaches should be optimized. The method proposed in this study will be investigated further by combining it with other methods using cameras, microphones, PC logs, etc.

6 Conclusion

Recognizing the state of a worker at a remote location is necessary for developing a sense of work sharing in remote collaboration. In this study, a worker state classifier using an ambient sensor method was developed by combining vibration sensor data and distance sensor data. We tested the approach for classifying the four states of a worker (‘writing on a desk,’ ‘typing using the keyboard,’ ‘viewing a PC monitor,’ and ‘leaving a seat.’) The results showed that the ambient sensor method can classify the four states with an accuracy of 66%. In future, the prototype module will be applied to remote collaboration to evaluate its effect on a local worker’s recognition of the remote worker.

7 Acknowledgments

We would like to thank Editage (www.editage.jp) for English language editing.

REFERENCES

- [1] P. Dourish and V. Bellotti, “Awareness and coordination in shared workspaces,” in *Proceedings of the 1992 ACM Conference on Computer-supported Cooperative Work, CSCW ’92*, pp. 107–114, 1992.
- [2] T. Rodden, “Populating the application: A model of awareness for cooperative applications,” in *Proceedings of the 1996 ACM Conference on Computer Supported Cooperative Work, CSCW ’96*, pp. 87–96, 1996.
- [3] C. Gutwin and S. Greenberg, “Design for individuals, design for groups: Tradeoffs between power and workspace awareness,” in *Proceedings of the 1998 ACM Conference on Computer Supported Cooperative Work, CSCW ’98*, pp. 207–216, 1998.
- [4] C. Gutwin and S. Greenberg, “A descriptive framework of workspace awareness for real-time groupware,” *Comput. Supported Coop. Work*, vol. 11, pp. 411–446, Nov. 2002.
- [5] L. S. Kennedy and D. P. W. Ellis, “Pitch-based emphasis detection for characterization of meeting recordings,” in *2003 IEEE Workshop on Automatic Speech Recognition and Understanding (IEEE Cat. No.03EX721)*, pp. 243–248, Nov 2003.
- [6] K. Otsuka, “Multimodal conversation scene analysis for understanding people’s communicative behaviors in face-to-face meetings,” in *Proceedings of the 1st International Conference on Human Interface and the Management of Information: Interacting with Information - Volume Part II, HCI’11*, (Berlin, Heidelberg), pp. 171–179, 2011.
- [7] K. Murao and T. Terada, “A combined-activity recognition method with accelerometers,” *Journal of Information Processing*, vol. 24, no. 3, pp. 512–521, 2016.
- [8] S. HASHIMOTO, T. TANAKA, K. AOKI, and K. FUJITA, “Improvement of interruptibility estimation during pc work by reflecting conversation status,” *IEICE Transactions on Information and Systems*, vol. E97.D, no. 12, pp. 3171–3180, 2014.
- [9] K. Iso, S. Ozawa, M. Date, H. Takada, Y. Andoh, and N. Matsuura, “Video conference 3d display that fuses images to replicate gaze direction,” *Journal of Display Technology*, vol. 8, pp. 511–520, Sept 2012.
- [10] A. Jones, M. Lang, G. Fyffe, X. Yu, J. Busch, I. McDowall, M. Bolas, and P. Debevec, “Achieving eye contact in a one-to-many 3d video teleconferencing system,” *ACM Trans. Graph.*, vol. 28, pp. 64:1–64:8, July 2009.
- [11] I. Feldmann, W. Waizenegger, N. Atzpadin, and O. Schreer, “Real-time depth estimation for immersive 3d videoconferencing,” in *2010 3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video*, pp. 1–4, June 2010.

Smart Robot based on Web Resource with Friendly Communication and Its Application to Library Guide

Yoichi Yamazaki*, Takuya Shinkai*, and Masao Isshiki*

*Dept. of Home Electronics, Kanagawa Institute of Technology, Japan
yamazaki@he.kanagawa-it.ac.jp

Abstract - Smart robot using web resource is proposed as a casual interface in smart house and smart society, where a robot interactively assists users utilizing existing web service in natural way (e.g., speech, gesture, etc.). In addition, a robot for casual service need to behave like a human beings in daily situation. Humor expression for a service robot is imprinted, where a robot interacts socially with users. We apply the proposed smart robot to the library guidance system. In the system, the robot gives users four kinds of guidance based on the existing web service of the library with two types of humor expression (play and support). The system is evaluated with subjective assessments, a face scale, and a smile expression time (per minute) for 30 subjects. The proposed smart robot shows a 264.7% increase compared to the situation without humor expression in the smile expression time. The proposed smart robot is able to realize a casual interface for IT services and IoT services including HEMS applications in smart house.

Keywords: Smart robot, Librarian robot, Cloud service.

1 INTRODUCTION

The ICT is evolving, and the IoT is spreading rapidly. In the light of the current situation, science and technology policy is carried out in the world, e.g., “Industry 4.0” in Germany, “Advanced Manufacturing Partnership” in the United States, and “Super Smart Society (Society 5.0)” in Japan. Super smart society (Society 5.0) is a society that is able to respond precisely to a wide variety of social needs, and a society in which all kinds of people can readily obtain high-quality services [1].

In a super smart society, things-driven service is provided by using the IT and the IoT resources, where networked things around human provide service automatically. In addition, a way of providing service become more and more important to a things-driven service in a daily situation. A robot interface is one of solution for this, and a robot for casual service need to behave like a human beings.

Smart robot using web resource is proposed as a casual interface at smart house in super smart society, where a robot interactively assists users utilizing existing web service in natural way (e.g., speech, gesture, etc.). Humor expression for a service robot is imprinted, where a robot interacts socially with users. We apply the proposed smart robot to the library guidance system. The system is evaluated with subjective assessments, a face scale, and a smile expression time for 30 subjects

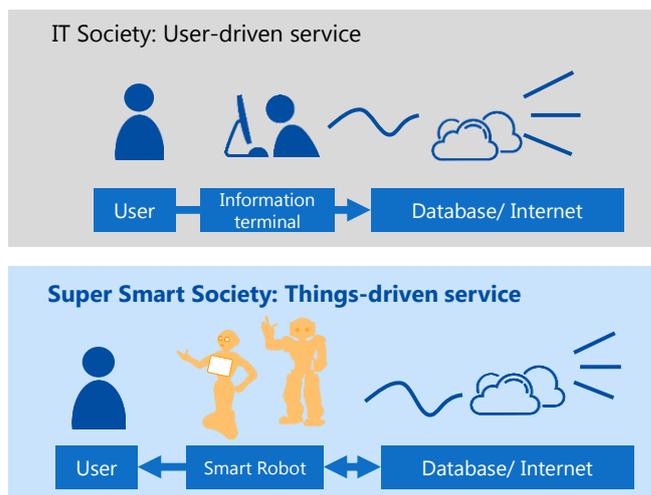


Figure 1: Super Smart Society and Things-driven service.

Concept of the proposed smart robot and humor expression are given in 2. The constructed smart robot system for library and its humor expression are described in 3. The experiments for the system are presented in 4.

2 SMART ROBOT IN SUPER SMART SOCIETY

The IoT is spreading rapidly, and society shift from the IT society to the super smart society, where the things connected the internet are utilized automatically and effectively. In the IT society, human uses IT resources through information terminals. This could be called user-driven service. On the other hand, things around human e.g. networked robots, smart phone, appliances, sensors, etc. use IT resources, and automatically offer their services to users with casual way in the super smart society. We define this type of service as things-driven service (shown as Figure 1), and propose smart robot system as a casual interface in the super smart society.

In this chapter, we define a smart robot first, and then show casual communication that is needed for smart robot.

2.1 Concept of Smart Robot

In the super smart society, smart robots assist human users on behalf of complex IoT systems surrounding. We define a smart robot as a robot that has following two conditions:

1. A networked robot (or thing) that can utilize existing IT resources for human users e.g. web contents in addition to IoT resources. This means that the robot need not only to connect

to the Internet as an IoT system component, but also to use existing web resource e.g., markup language document etc..

2. A robot has abilities for casual communication like human. These abilities basically includes speech and nonverbal communication. The robot for casual communication has been proposed in [2, 3], and humor expression explained in the next section is taken as developmental ability for casual service communication.

2.2 Humor Expression for Casual Service Communication of Robot

Robots which provide service to human user in daily situation need to adopt a good attitude toward helping and taking care of the customer. To build good attitude, a robot is required to take friendly behavior and to respond appropriately. As for friendly behavior, emotional interaction enable robots to be accepted favorably [2-4]. As for appropriate response, a robot requires a high-performance speech recognition and a function to recover from a recognition error. In human communication, humor works effectively in the operation [5], and humor has possibilities to make up for a mistake in speech recognition of robots.

A humor expression is proposed for communications robot with speech recognition. Humor is fall into following three fundamental classes i.e. 1) ludic humor that entertains others, 2) offensive humor that attacks others to obtain a sense of superiority, and 3) supportive humor that encourages and forgives others [6]. Ludic humor and supportive humor are used as proposed humor expression to decrease mental stress by error in human-robot communication.

3 SMART ROBOT AND ITS HUMORE EXPRESSION FOR TINGS-DRIVEN SERVICE

As an application of the proposed smart robot, a guide robot is constructed for a library. In this chapter, the hardware configuration, the application components and its humor expressions for library system is described.

3.1 Smart Robot System for a Library in a University

A smart robot system for a library in a university is constructed as a demonstration of a smart robot application, where a companion robot NAO assists users by using existing library guide contents on the website. The system flow is shown in Figure 2. We prepare two types of smart robot system. One is a system for public libraries, where the system is connected to the internet. In this case, the robot send requests to the display server according to conversation with users. The server searches results from the database on the web, and outputs the result on the display connected the server PC.

The other is a system for smart house environment, where the system can use HEMS communication based on ECHONET Lite. In this case, the robot send requests to the ECHONET Lite-ready TV directly, and the TV searches and shows the result on its screen.

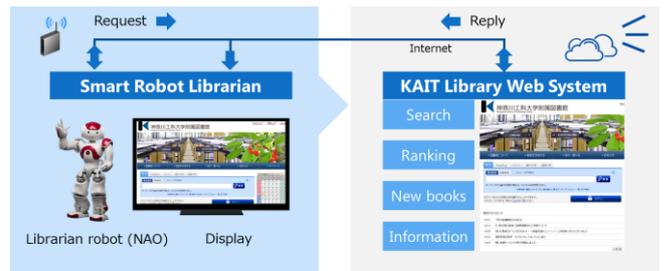


Figure 2: Implemented smart robot system for library guide.



Figure 3: The main contents of the library's website. This page is originally in Japanese, and translated in English by Google Translate.

These two systems work the same way from the viewpoint of users. At the moment, we use the former system for substantive experiments shown in 4, for public libraries generally don't have many ECHONET Lite-ready appliances at the moment. In the days ahead, smart house environment are spreading rapidly and widely in both of private and public space.

3.2 Smart Robot System with Humor Expression for a Library in a University

The proposed system is applied to the Kanagawa Institute of Technology library (KAIT library). The main contents of the library's website are shown in Figure 3, where mainly following four functions are prepared, i.e. 1) searching book, 2) monthly rental ranking, 3) new arrival books, and 4) latest information.

These functions are used by using the smart robot based on speech interaction. Users speak to the robot, and the robot creates URL query parameter based on the speech. And then the robot connects the website with URL query parameter, and gets the information that the users want from the result page. The obtained information is presented to the users by its speech and the screen display.

Two types of humor (ludic humor and supportive humor) are used as humor expression in the system. At the beginning of the guidance, the robot introduce itself with ludic humor aiming to break the tension. In the speech recognition, the robot occasionally expresses ludic humor. The robot

expresses supportive humor when the robot fail to recognize words from users.

In the experiment scenario described in 4, ‘monthly rental ranking’ and ‘latest information’ are used as the website service, and ‘how to use the service of book’ is additionally prepared as general service for NAO, where two types of expression (without/with expressions) are performed. The examples of the scenes in the scenario are as follows: (The italic and underlined parts are humor expressions. The single underline shows ludic humor parts. The double underline shows a supportive humor part in the sentence.)

Interaction A): Interaction without humor (Control).

Robot: Hello. Welcome to the Kanagawa Institute of Technology library.

I am NAO.

I will show you ‘How to use the service of book,’ ‘Monthly rental ranking,’ and ‘Latest information’.
What do you want to know?

User: ‘Latest information.’

Robot: (When the robot fail to recognize words from users)
Excuse me, but I cannot hear you. Could you please tell me what you want to know again? I will show you ‘How to use the service of book,’ ‘Monthly rental ranking,’ and ‘Latest information’. What do you want to know?

Interaction B): Interaction with humor (Proposed).

Robot: Hello. Welcome to the Kanagawa Institute of Technology library.

I am NAO. *Call me plain Nao-chi*! I want to have an orange in the kotatsu** on a cold day. Oh! Where is the orange oranged***?*

I will show you ‘How to use the service of book,’ ‘Monthly rental ranking,’ and ‘Latest information’.
What do you want to know?

User: ‘Latest information.’

Robot: (When the robot fail to recognize words from users)
Excuse Medium-rare diced steak? ****. I’m too focus on your beautiful voice. Could you please tell me what you want to know again? I will show you ‘How to use the service of book,’ ‘Monthly rental ranking,’ and ‘Latest information’. What do you want to know?

* ‘-chi’ is one of popular ending of a word as nickname in Japanese.

** Kotatsu is a small table with an electric heater underneath and covered by a quilt. A kotatsu is one of popular heating appliances in Japan. Mandarin orange is often eaten in a kotatsu.

*** A pun using the words ‘orange’ and ‘arrange’. In fact the experiment is conducted in Japanese, and the real sentence in Japanese is as follows:
‘ARE? Mikan ga Mikkan-nai!’ (This is same meaning as the English expression described above.)

**** A pun using the words ‘me’ and ‘medium’. the real sentence in Japanese is as follows:

‘GomennaSAIKORO Steeki!’ (The meaning is ‘Excuse Me-Diced steak?’) This is a pun using ‘Gomennasai (Excuse me/Sorry)’ and ‘Saikoro (dice)’ in Japanese.)

Effects on users by these humor expression are evaluated in subjective ways and an objective way. In subjective ways, users are directly offered two questionnaire, i.e. a subjective assessment and a stress levels assessment with a face scale [7] to evaluate their feelings and moods. In objective evaluation, smile expression time is measured implicitly. The users are left uninformed of the measuring of smile. Smile– with or without awareness of expression– has a positive impact on stress recovery [8]. Therefore, we regard the smile expression time as an index of stress in communication. In other words, stress is released when a user is smiling.

4 SYSTEM EVALUATION OF SMART ROBOT FOR LIBRARY

The proposed smart robot system is evaluated by 30 users explicitly and implicitly. The experiment is conducted at main floor in the Kanagawa Institute of Technology library. Two conditions of the robot are prepared in the experiment. One is a robot that interacts with users without humor expressions (Control). The other is a robot that interacts with humor expressions (Proposed). The specification of the experiment is as follows.

4.1 Evaluations Items

In the experimental situation, the robot guidance is given to users according to request. In the interaction, users evaluate subjective assessments and their stress level in a face scale. At the same time, smile expression time is implicitly measured based on dynamic image analysis.

4.1.1. Subjective Assessments Questionnaire for Robot Impression

In the subjective assessments, the following six items are evaluated on a scale of one to five.

Item 1 and Item 2 are related to fullness in conversation.

Item 1: Length of sensation time spent in conversation. The length is assigned from ‘short’ to ‘long’ with a score of one to five.

Item 2: Pleasant in conversation. The pleasant is assigned from ‘unpleasant’ to ‘pleasant’ with a score of one to five.

Item 3 and Item 4 are related to familiarity of the robot.

Item 3: Companionable of the robot. The familiarity is assigned from ‘uncompanionable’ to ‘companionable’ with a score of one to five.

Item 4: Humor in conversation. The pleasant is assigned from ‘unfunny’ to ‘funny’ with a score of one to five.

Item 5 and Item 6 are related to stress in conversation.

Item 5: Response in conversation. The response is assigned from ‘bad’ to ‘good’ with a score of one to five.

Item 6: Stress in conversation. The stress is assigned from ‘stressful’ to ‘stress-free’ with a score of one to five.

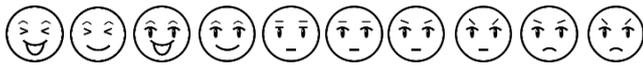


Figure 4: Face scale of one to ten used in the experiment.

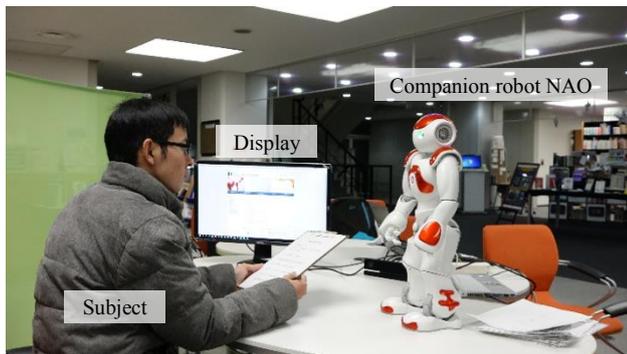


Figure 5: Experimental situation for evaluation of the proposed system in KAIT library.

4.1.2. Stress Level Assessments with Face Scale Measure

Stress level is evaluated with a face scale in the conversation with the smart robot. A face scale enable to check a feeling of users. The face scale of one to ten used in the experiment is shown in Figure 4. The Face scale shows the faces above range from very happy (its score is 1) to very stressful (its score is 10). Users check the face which is best suited to their feeling.

4.1.3. Smile Expression Time

Both Subjective assessment questionnaire and face scale measure are subjective evaluation of users. As an objective assessment data of familiarity of the system, we measure an expression time when users smile on the face based on the recorded movie. As an indication of familiarity, we define a smile expression time as follows:

Smile Expression Time per minutes

$$= \frac{\text{Smile Expression Time in interaction}}{\text{Total Time of interaction}} \times 60 \quad (1).$$

Users are explained the experimental scene is recorded, and are not explained the smile expression time is measured.

4.2 Experimental Procedures

The procedures for the experiment is as follows:

- Step 1:** The operator informs an experimental participant about the smart robot system, the subjective assessments, and the stress level evaluation with the face scale. The operator explains to the participant that the experimental scene is recorded, and does not explain that smile expression time is measured.
- Step 2:** The operator chooses the robot’s expression (with humor or without humor expression). The choice is selected randomly. The robot introduces itself to a user with the selected expression.
- Step 3:** The user freely chooses one from the three guidance (1. Lending service of book, 2. Monthly ranking, and

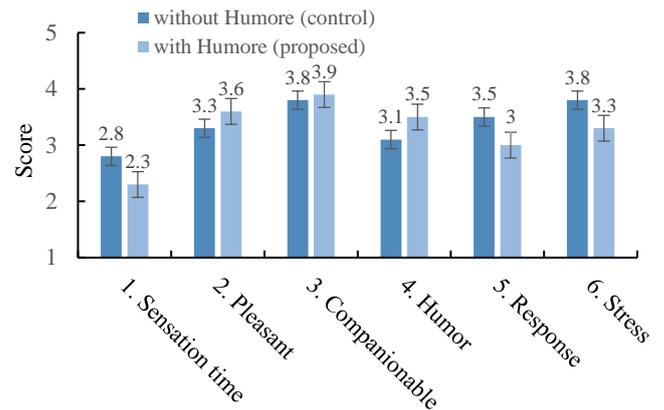


Figure 6: Result of subjective questionnaire (1 to 5).

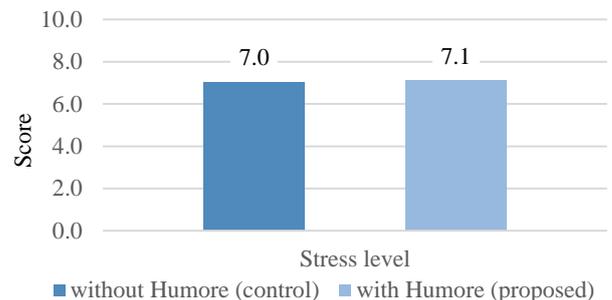


Figure 7: Result of stress level assessment with the face scale (1 to 10).

3. Information). When the user chooses ‘2. Monthly ranking’, the robot make sure to fail recognition the word. This is an intentional stressor in conversation to evaluate stress level of the user.

- Step 4:** The user repeats Step 3 on all three guidance.
- Step 5:** The user answers the subjective questionnaire and the stress level check with the face scale (shown in Figure 4).
- Step 6:** Step 2, Step 3, Step 4, and Step 5 are repeated on the other type of robot expression (with humor or without humor expression).

The experimental situation is shown as Figure 5. In the experiment, several humor expressions are prepared, and are changed as situation demands.

4.3 Results

The results of subjective questionnaire are shown in Figure 6. The proposed robot with the humor expression has shown more positive tendency as for ‘1. Sensation’ time, ‘2. Pleasant’, ‘3. Companionable’, and ‘4. Humor’, and has shown more negative in ‘5. Response’ and ‘6. Stress.’

The results of stress level assessment with the face scale are shown in Figure 7. There is little difference between the control robot without the humor expression and the proposed robot with the humor expression.

The result of the smile expression time is shown in Figure 8. The results have shows the proposed smart robot obtain a 264.7% increase compared to the situation without humor expression in the smile expression time per minutes.

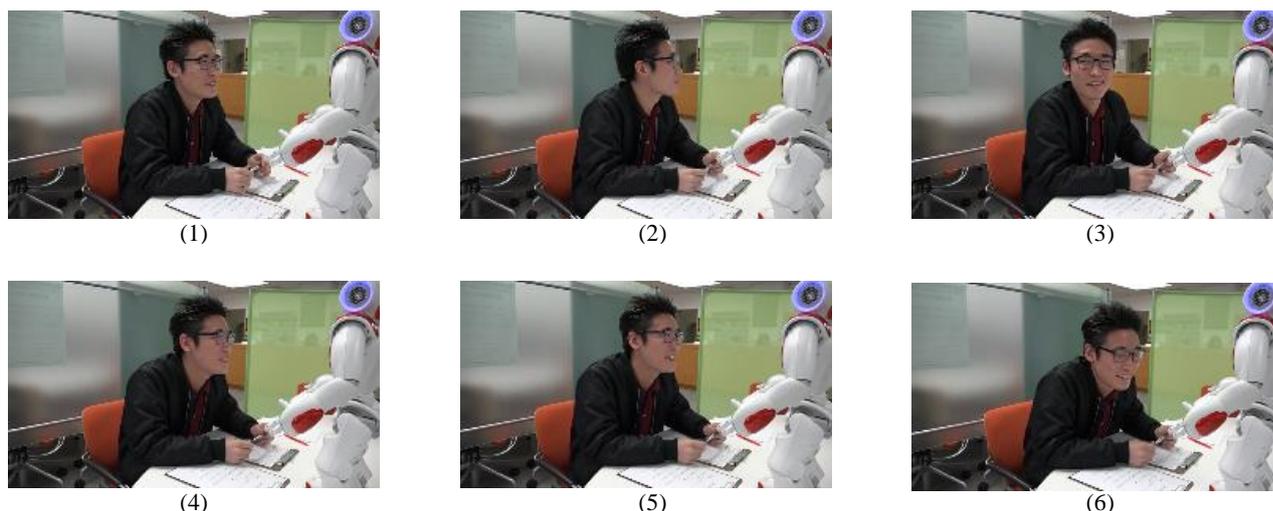


Figure 9: User reaction to the humor expression (one frame per second).The user has shown ‘a little stressful (2)’ in the subjective questionnaire.

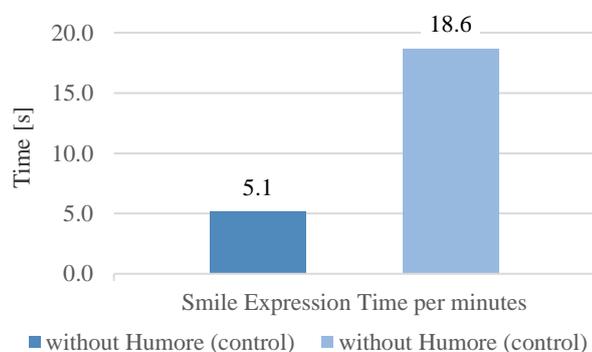


Figure 8: Smile expression time per minutes.

4.4 Discussion

The result of the smile expression time has shown the most marked difference between the proposed robot and the control robot in the three result. These all results have shown the users express simile unconsciously, though they feel a little stress in the interaction consciously by the humor expression. Figure 9 shows the user has been smiling in the interaction with the proposed robot, though he has felt a little stress in the interaction (He has shown ‘a little stressful (2)’ in the subjective questionnaire.) Human reduce their stress by just smiling on the face in general. Therefore, we conclude the results shows the proposed expression has positive effect for the service robot with speech recognition function.

5 CONCLUSION

Smart robot using web resource and its humor expression has been proposed as a casual interface in smart society. We have applied the proposed smart robot to the library guidance system. The three typed evaluation has been conducted for 30 subjects, and the proposed smart robot has shown a 264.7% increase compared to the situation without humor expression in the smile expression time. The system is able to be widely applied to existing IT system in daily life.

ACKNOWLEDGMENTS

The smart robot system in this research has been performed in the Kanagawa Institute of Technology library (KAIT library). We would like to thanks to Satoshi Watanabe and acknowledges to the KAIT library.

REFERENCES

- [1] The 5th Science and Technology Basic Plan, Government of Japan (2016).
- [2] Y. Yamazaki, et al., Mascot Robot System by integrating Eye Robot and Speech Recognition using RT Middleware and its Casual Information Recommendation, Proc. of the 3rd International Symposium on Computational Intelligence and Industrial Applications, pp.375-384 (2008).
- [3] Y. Tang, et al., Deep Level Situation Understanding for Casual Communication in Humans-Robots Interaction, International Journal of Fuzzy Logic and Intelligent Systems Vol. 15, No. 1, pp. 1-11 (2015).
- [4] Cynthia Breazeal, Emotion and sociable humanoid robots, Int. J. Human-Computer Studies vol.59, pp.119-155 (2003).
- [5] J. Maruyama and K. Fuji, The effect of humor in the workplace on mental/physical health and self-evaluation of job performance, The Japanese Journal of Psychology, vol.87, No.1, pp.21-31 (2016).
- [6] Y. Ueno, Studies of Humor phenomenon and Classification of Humor, Japanese Journal of Social Psychology, vol.7, No.2, pp.112-120 (1992).
- [7] K. WADA, et al., Robot Therapy at a Health Service Facility for the Aged : Evaluation of Psychological Effects of One Year Experiment, Transactions of the Society of Instrument and Control Engineers Vol.48, No.4, pp.386-392 (2006).
- [8] Tara L. Kraft, Sarah D. Pressman, Grin and Bear It: The Influence of Manipulated Facial Expression on the Stress Response, Psychological Science, Vol.23, No.11, pp.1372-1378 (2012).

A Comparative Study in Enterprise Systems Dealing with Images: MongoDB vs. MySQL

Tsukasa Kudo[†], and Yuki Furukawa[†]

[†]Faculty of Informatics, Shizuoka Institute of Science and Technology, Japan
kudo.tsukasa@sist.ac.jp

Abstract - Nowadays, it is possible to enter various kinds of data such as the images and videos into systems with the development of various sensor such as surveillance cameras, wearable devices and so on. To manipulate these data, which is difficult to be dealt with the relational databases in terms of capacity and performance, various NoSQL databases have been put to practical use. Especially, MongoDB provides GridFS interface for such a data manipulation, and we confirmed that it was effective in our previous study. On the other hand, it is pointed out: since some data manipulations such as the join operation are not provided in NoSQL databases, there are problems to migrate the system using the relational database to that using NoSQL database. In this study, we tried to migrate to MongoDB the data manipulation patterns, which is extracted from the actual enterprise system constructed by the relational database. Then, performed the comparative evaluations on them about the product volume and performance. As a results we found that this migration was possible, though the program volume tended to become large; there was the extremely large difference between MongoDB and MySQL depending on the type of data manipulation.

Keywords: MongoDB, database, GridFS, join operation, production management system

1 INTRODUCTION

Nowadays, various devices are spreading rapidly, such as smartphones, surveillance cameras or various wearable devices. As a result, it is becoming possible to enter various data efficiently into the systems using such an inexpensive entry device [3], [5]. Therefore, even as for the enterprise systems, it is expected that the system can be operated more efficiently by utilizing not only conventional character and numeric data but also various kinds of sensing data such as images and videos.

In order to store and manipulate such a data, various NoSQL databases have been proposed and put to practical use [12]. Among them, there is MongoDB [2] which is a kind of the document-oriented NoSQL database: it stores the data as documents of semi-structured data model expressed by JSON; it equips GridFS interface in particular to treat the enormous data efficiently [11]. On the other hand, it have been pointed out that there were some problems to apply it to the enterprise systems: it does not maintain the ACID properties of the transaction; it does not equip the operation for the plural tables like the join operation of the relational databases (RDB) [13].

Here, we are advancing the study on the production management system utilizing images and videos in order to improve the efficiency of the inventory management works. That is, conventionally, the inventory quantity of many kinds of parts must be counted and it makes the workload higher. So, in our previous study, we had conceived the method, in which the manager judge visually whether there was the necessary inventory quantity by using the images and videos [7]. In addition, this work can be performed in the office based on the inventory plan calculated in advance, instead of counting in the field.

Here, though the target production management system uses MySQL for the database, we have shown that the enormous data such as video can be manipulated much more efficiently by MongoDB than MySQL [8]. So, we performed the survey as the feasibility study to migrate from MySQL to MongoDB for this system. And, we found some comparative studies between MongoDB and MySQL have been performed [1], [4]. However, some problems such as the transaction processing and join operation, was pointed out to apply MongoDB to the enterprise systems, so we could not find the study on the actual enterprise system. For these problems, we have already composed the transaction feature for MongoDB to maintain the ACID properties in our previous study [6].

So, our goal of this study is to confirm the feasibility of applying MongoDB to the target enterprise system, in which the above-mentioned data manipulations such as the join operation are used. In this paper, firstly, we extract the data manipulation patterns used in this system, and show that these manipulation can be migrated to MongoDB. Then we show the comparative evaluation of MongoDB and MySQL about the product volume and performance.

The remainder of this paper is organized as follows. Section 2 shows the abstract of the target system to clarify the precondition. Section 3 shows the correspondence of data manipulation used in the system between MongoDB and MySQL, then we show the implementation using MongoDB. Section 4 shows the results of comparative evaluations between MongoDB and MySQL, and we discuss on the result in Section 5. Lastly, Section 6 concludes this paper.

2 TARGET SYSTEM

2.1 Target Function of Production Management System

The target enterprise system of this study is an actual production management system of some company. And, some of their function have been already in operation; the others are

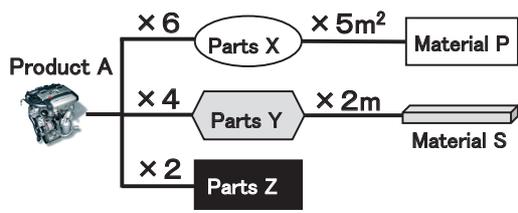


Figure 1: Structure of BOM of Target System.

Spec_id	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
001A				1			1				1	
002A	4	5	4	2	1	4	5	3	2	2	2	
002B						2	1			3		
001A	4	2	2	3	2	7	4	4	4	4	2	1
003H	1		3	1	5	2	2					
109H	4	2	2	3	2	7	4	4	4	4	2	1

Figure 2: Monthly total of each spec

currently under development. We use MySQL for the relational database management system (RDBMS), and the calculation processing of each function is executed collectively by batch processing, then the results are stored into the database. We use Excel to entry the source data or to output the results. We show the outline of the target system below.

The first is the material requirement calculation function, which is already in operation and manages the bill of material (BOM) [14] as shown in Fig. 1. In this figure, product A consists of 6, 4, 2 parts X, Y, Z respectively, and parts X, Y is manufactured from 5m² board material P and 2m stick material S respectively. For parts Z, the commercial goods are purchased. In this way, by managing the BOM, it is possible to calculate the material cost of A based on the price per unit of P, S and the price of Z.

The configuration of this processing is as follows. The data for the calculation is composed of the BOM, products, parts, materials as shown in Fig. 1. And, they are stored in the tables, and it is changed if necessary by MySQL for Excel which is a linkage tool between MySQL and Excel. Then, the calculation processing is executed for all the data collectively, and there is not necessary to specify the parameters. So, its process is described only by SQL statement, and they are executed as a batch file of Windows. Lastly, through the view table, the calculation results are converted to the various forms to be handled easily, and output by above-mentioned MySQL for Excel.

The second is the production planning function, and the

		Mar.							Apr.				
Order_id	Spec_id	22	23	24	27	28	29	30	31	1	3	4	5
		Wed	Thu	Fri	Mon	Tue	Wed	Thu	Fri	Sat	Mon	Tue	Wed
AP000001-01	001A	▼	●					■			▲		
AP000002-01	002A		▼	●					■			▲	
AP000003-01	003H		▼	●					■			▲	
AP000002-02	002A		▼	●					■			▲	
AP000004-01	104A		▼	●					■			▲	

Remarks) ▼: Manufactured; ●: Start to prepare shipment; ■: Delivery; ▲: Used

Figure 3: Manufacturing schedule

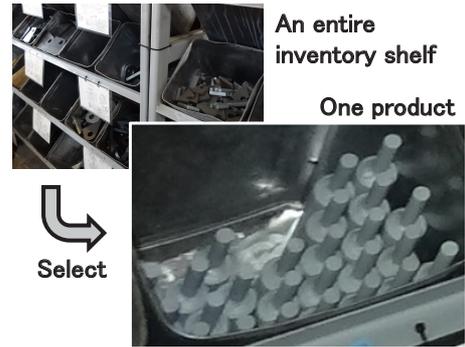


Figure 4: Inventory management utilizing images

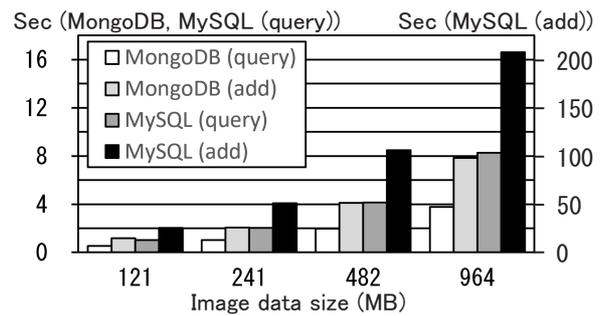


Figure 5: Evaluation result of performance comparison

plan is made based on the received or expected receipt orders. The contents of the order are specified by the specification (spec) sheet composed of each product type and its quantity, which has spec identifier (ID). And, this system targets the common products specified by the spec ID. That is, though the order includes the custom ordered products which are individually specified at each time, they are not managed by this system. We show the sample of the output documents in Fig. 2 and 3. Fig. 2 shows the monthly total number of each spec, and it is used to grasp the long-term order status. Fig. 3 shows the monthly work plan which is made per order, and it is used to grasp the daily mile stone. Fig. 3 shows the monthly work plan, which is made per order and used to grasp the daily mile stone. We are currently conducting the operation test of this function and preparing the necessary data.

In this function, since the parameter must be specified such as the target month, we composed this by Excel and Excel VBA (Visual Basic for Applications). That is, parameters are entered from Excel sheet, then Excel VBA makes SQL statement then start its batch file. Lastly, its results are processed to output forms by Excel VBA similar to the first. Moreover, we composed some MySQL data manipulations by the stored procedure or stored function [9]. For example, to make the work plan, we must calculate the number of business days excluding holidays. So, it is necessary to make the calendar to show the number of each business day from the beginning of the year. And, since such processing includes the iterative processing, it cannot be made by only the simple SQL statements.

The third is the inventory management function that saves the status of each product shelf as the images and videos. Fig. 4 shows the inventory images, and these are stored into the

database. The aim of this function is to offer the inventory image with its necessary quantity information to the manager at the office to confirm the satisfaction of its inventory. So, it is necessary that the manager can set the inventory status based on the confirmation results, and the images of the specified products must be queried based on the manufacturing schedule shown in Fig. 3. To introduce this function, we are currently conducting the evaluation of its prototype. Moreover, we have already conducted the comparative evaluation on the performance between MongoDB and MySQL about the video data, and found MongoDB was much more efficient especially to insert the data as shown in Fig. 5. As for the insertion in this case, MongoDB was 25 times faster than MySQL [8].

Incidentally, the actual system is composed of various functions besides the above: it calculates the MRP (Material Requirement Planning) [14], which is the necessary quantity of the parts and materials, by the linkage of the production plan and BOM; the various management documents are output, and so on. However, since the data manipulation patterns are covered by the above-mentioned cases, we conduct the comparative study on them in this study.

2.2 Database Structure and Data Manipulations

Figure 5 shows the ER diagram of the target system; below, we show the tables and attributes of the ER diagram in italic. Here, we show only the main tables and attributes for the sake of simplify. (1) of Fig. 5 corresponds to the material requirement calculation function, and parts (*part*) and products (*product*) are associated by BOM (*BOM*). And, the following price is set to the unit price (*price_unit*) of parts: *price_unit* of *part_price* is set in the case where the part is purchased; price per 1kg (*price.kg*) of material is set in the case where the parts is manufactured from the material associated by material ID (*mat_id*).

(2) of Fig. 5 corresponds to the production planning function, and the data of *calendar* such as the number of business days is calculated by the holiday information (*holiday*); order and product are associated with specification, which is composed of *specification* and *spec*. The schedule data shown in Fig. reffig:schedule is calculated and saved into *manufacture_plan*, in which each mile stone date is included: manufacturing completion date (*m_date_num*), start date to prepare shipment (*c_date_num*), delivery date (*d_date_num*) and used date at the ordering company (*u_date_num*).

(3) of Fig. 5 corresponds to the inventory management function, and saves the inventory status of products as images and videos. *product_shelf* indicates the product stored on each product shelf. And *inventory* shows the status of the inventory: the image or video name (*img_name*) saved in *image*; the correspondence data between them and product shelf (*shelf_id*); the inventory quantity (*quantity*) if necessary. Here, since the images and videos are chaptered and save at any time, *img_name* is used as the primary key and the relationship between *product_shelf* and *inventory* is many-to-one. Firstly, the images and videos data is saved from the camera into the folder of the PC, then saved into

```
SELECT * FROM part_price AS a
WHERE a.est_ymd = (SELECT
MAX(b.est_ymd) FROM part_price AS b
WHERE a.part_id = b.part_id);
```

Figure 7: Max value query by self-join

image; to use them, the necessary data is downloaded to the folder of the PC.

We extracted the data manipulation patterns for these tables from the functions mentioned in Section 2.1, and got the patterns shown below. Incidentally, the basic CRUD data manipulations on the single table are excluded.

- (a) Join operation: in (1), each part is jointed with the products which use it respectively, and the results are saved into *parts_cost*. So, the join operation between *BOM* and *parts* is performed.
- (b) Iterative operation: in (2), to set the number of business day (*date_num*) of *calendar*, the division of the business day and holiday are set to the column *holiday* of *calendar* firstly. Then, the numbers of business date are set sequentially from January 1st, that is, it constitutes the iterative operation. Incidentally, it is implemented by the stored procedure in MySQL.
- (c) Grouped aggregation operation: in (1), the product of material cost and quantity, which are expressed by *cost* and *p_quantity* of *parts_cost* and calculated in (a), are aggregated for each product, and stored into *product_price*.
- (d) Selection of record with self-join operation: in (1), since *part_price* has a history on the estimate date (*est_ymd*), the record having the max estimate date must be queried for each *part_id*. In the SQL statement, this is expressed by the sub query with self-join operation as shown in Fig. 7.
- (e) Images and videos operation: in (3), the images and videos of the inventory shelves are stored into *image*, so these data must be inserted and queried. In MySQL, this is executed by “load_file” function to insert, and “select into dumpfile” statement to query.

3 IMPLEMENTATION OF DATA MANIPULATION USING MONGODB

3.1 Implementation policy

In MongoDB, the mongo shell is provided for interactive data manipulation, which has JavaScript interface. And, similar to the SQL statements, JavaScript files can be executed as the batch file, or as a function like the stored procedure of SQL. In this study, we implemented the mongo shell as batch file of Windows as shown in Fig. 7. In Fig. 7, “JSfile.js” is the JavaScript file including the mongo shell methods; and, it is executed by inputting to “mong” command; then the execution results are output to “out.csv” file by print statement of JavaScript.

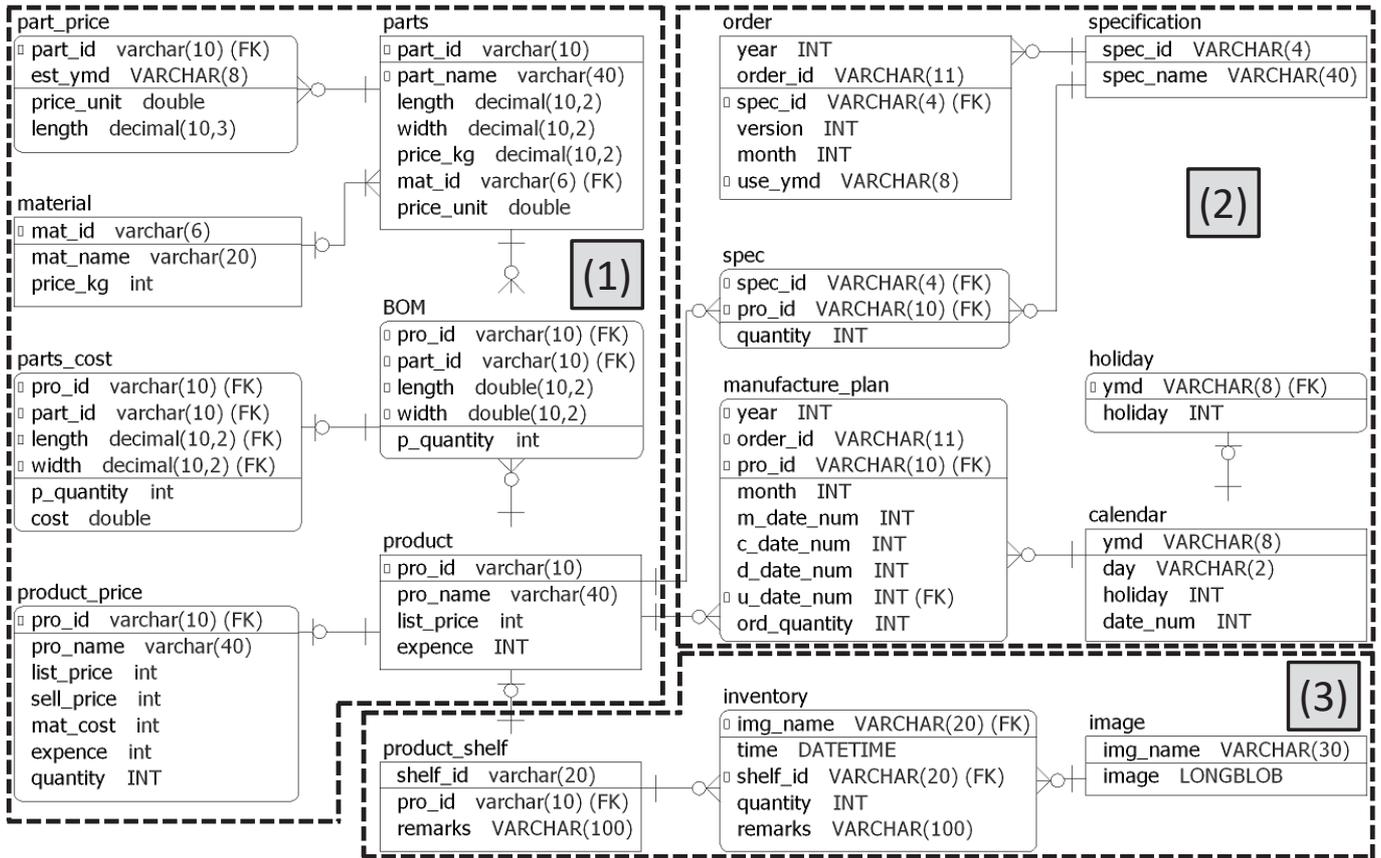


Figure 6: ER Diagram of database

```
> mongo < JSfile.js > out.csv
```

Figure 8: Batch file of mongo shell

Below, we show the implementation of each operation mentioned in Section 2.2. Incidentally, we describe only the main attributes and operations for the sake of simplicity. In the actual system, related attributes and operations are added to the following logic.

3.2 Implementation of join and iterative operation

In the mongo shell, the join operation is not provided. So, we linked the collections as follows: firstly, we copied collection corresponding to “many” of many-to-one to the temporal new collection; then, we added the fields of the collection corresponding to “one” to the above collection. In this way, we could create the result collection of the join operation.

In Fig. 9, we show the case of (a) in Section 2.2, in which parts and BOM are joined to create parts_cost. At (1) of Fig. 9, parts_cost is created by copying BOM, then cost field of parts_cost is set price_unit field of parts at (2) as follows. Firstly, find method at (3) corresponds to select statement of SQL, and query all the documents of parts. Here, the documents are sequentially set to partsRec as same as the cursor operation of SQL. Next, update method at (4)

```
// (1) copy BOM to parts_cost
db.BOM.copyTo("parts_cost");
// (2) update parts_cost to join parts
var partRec; // variable (document of parts)
var part=db.parts.find(); // (3) find method
while(part.hasNext()){
  partRec=part.next(); // get next document
  db.parts_cost.update( // (4) update method
    {zairyou_id:partRec.zairyou_id,...
      haba:partRec.width}, // query condition
    {$set:{cost:part.price_unit}}, // set cost
    {multi:true}); // update multi documents
}
```

Figure 9: Join operation procedure of MongoDB

which corresponds to update statement of SQL updates the value of cost field of all the document that matches the query condition shown by the first parenthesis “{ }”, which expresses the pairs as of “{field name:variable}”. And, the query condition here is that all the field values are equal to these variable values. Incidentally, though parts_cost collection does not have cost field, it is added according to the setting {multi : true}.

Incidentally, in the case where the join operation is performed on only a part of documents matching the specified

```

// (1) total cost per product_id
Var mat_cost=db.parts_cost.aggregate
  ({ $group: { _id: { product_id: "$product_id" },
    cost: { $sum: "$cost" } } });

// (2) document with latest date
var part_price=db.part_price.aggregate
  ({ $match: { parts_id: "P0001" },
    $group: { _id: "$part_id",
      latest: { $max: { est_ymd: "$est_ymd" } } } });
var partRec=part_price.next();
var laRec=db.part_price.findOne
  ({ parts_id: "P0001", est_ymd: laRec.latest });

```

Figure 10: Aggregation method of MongoDB

query condition, instead of the above-mentioned case for all the document, only the target documents are inserted at (1). For this operation, insert method corresponding to INSERT statement of SQL is used.

Next, the iterative statement such as while or for can be used in the mongo shell. So, we implemented the iterative operation to create *calendar* shown in (b) of Section 2.2 by using these statements, like the stored procedure of SQL.

3.3 Implementation of aggregation and seff-join operation

In the mongo shell, aggregate method is provided, which corresponds to the aggregation operator and group by expression of SQL. So, as for the aggregation operation of material cost for each parts in (c) of Section 2.2, it can be executed by this method. In this method, as shown in (1) of Fig. 10, \$group expression shows the fields to be aggregated, and \$sum expression shows the aggregation method like the SQL statement. Incidentally, the aggregation results can be get by the cursor operation like Fig. 9.

Similarly, as shown in (2) of Fig. 10, selection operation of record having max value shown in (d) of Section 2.2 can be performed by aggregate method, in which the latest estimated date is queried from *parts_price*. Here, since MongoDB does not provide the join operation, we configured to query the target document again from *part_price* using the queried estimated date and *parts_id* value. In Fig. 10, \$match expression in aggregate method specifies the query condition. Also, findOne method queries only the single document, and in this figure, it queries as of the query condition that *parts_id* is "P0001" and *est_ymd* is the queried estimated date.

3.4 Images and Videos operation

The upper limit of the document size of MongoDB is 16 MB, and the GridFS interface is provided for data exceeding this limit. Using this interface, the image and video data is saved into GridFS collection divided from other attributes. And, since the data insertion and query are performed by using mongofiles command not the mongo shell, we configured

```

REM (1) insert operation of image from file
mongofiles -d iwin2017
  put 2017_5_A-1-3-1.JPG -l A-1-3-1.JPG

REM (2) query operation of image into file
mongofiles -d iwin2017
  get 2017_5_A1-1-1.JPG -l A-1-3-1.JPG

```

Figure 11: Image insertion and query command

Table 1: Comparison of CRUD operation

MySQL	MongoDB
SELECT	find(), findOne()
INSERT	insert()
UPDATE	update()
DELETE	remove()
Group BY clause	aggregate()
Stored procedure	JavaScript
Stored function	JavaScript
<i>(Join operation)</i>	
JOIN syntax	[many].copyTo() [one].find() (use cursor) [many].update()
<i>(Query record with max value)</i>	
self-JOIN operation + subquery	aggregate() findOne()
<i>(Image and video operation)</i>	
INSERT + LOAD_FILE()	MONGOFILES command (put)
SELECT INTO DUMPFIL	MONGOFILES command (get)

to perform this command by batch files which is separated from the JavaScript files.

We show the examples of these commands in Fig. 11. Here, "-d" indicates the database, and "-l" indicates the file name on the disk. That is, we can save the image with the different name from the name as of disk file.

Finally, in Table 1, we show the summary of the implementation method comparison between MySQL and MongoDB.

4 IMPLEMENTATIONS AND COMPARATIVE EVALUATIONS

In order to demonstrate the target production management system can be constructed by using MongoDB, we implemented the principal part of this system by using MongoDB according to the correspondence shown in Table 1. Then, we conducted the comparative evaluations about the program volume and execution performance between MongoDB and MySQL.

4.1 Implementation using MongoDB

First, as for the material requirement calculation function, we implemented the process to create *product_price*: the cost

Table 2: Comparison of program volume

	MySQL			MongoDB		
	SQL	Else	Total	Shell	Else	Total
(1) Material	15	0	15	24	29	53
(2) Plan (C)	8	64	72	12	91	103
(2) Plan (O)	11	7	18	8	24	32
Total	34	71	105	44	144	188
(2) Plan (P)	0	180	180	0	180	180

(1): Material requirement calculation function
(2): Production planning function
C: create data; O: output to csv file; P: print document

data of *parts_price* and *material* is reflect into *parts*, then *product_price* is created from *parts* via *parts_price*. We implemented this processing using the mongo shell as the batch file shown in Fig. 8.

Second, as for the production planning function, we implemented the processing to create *calendar* from *holiday*, and to make the csv files for the aggregation and schedule document, which is shown in Fig. 2 and Fig. 3 respectively. We implemented this processing using the above-mentioned batch file, and we embedded the parameters in the JavaScript program directly without linking with Excel for the sake of simplicity.

Third, as for the inventory management function, we implemented two processing shown in Fig. 4. Incidentally, these implementation method is same as MySQL except the execution command as shown below.

One is the processing to saves the pictures and videos of the product shelves into *image*, and to insert the correspondence data between *product_shelf* and *image* into *inventory*, that is, this is the correspondence between the shelves and the images or videos. To save the images and videos, the file name in the camera must be grasped to make the insertion program. So, we implemented this processing using Excel VBA to make the insertion batch file, in which insertion is executed by *mongofiles* command. Then, we implemented the correspondence data insertion program using the mongo shell, and we executed it using the batch file.

The other is the processing to down load the images and videos. Similar to above, we implemented this processing to executed by *mongofiles* command, which is made by using Excel VBA based on the given query condition: specified shelves, or the product shipment date and so on.

4.2 Comparative Evaluations of Program Volume

In order to evaluate the production efficiency comparatively in the case of implementing with MongoDB and MySQL, we counted the number of source lines of the programs respectively. Table 3 shows this results, and (1) shows the material calculation; (2) shows the production planning function. Incidentally, since the database access program for the inventory management function is made by using Excel VBA as mentioned in Section 4.1, we omit this evaluation. Here, (2) is

Table 3: Comparison of Elapsed time

Processing	MySQL	MongDB	Ratio	Manipulate
Material	1.07	11.90	11.1	(a), (c), (d)
Calendar	9.58	5.10	0.5	(b)
Plan	0.32	26.84	83.0	(a) (3-join)
Image (in)	23.48	260.78	11.1	(e)
Image (out)	0.24	77.76	324.0	(e)

Remarks: Ratio=MongoDB/MySQL

divided into three processing: Plan (C) shows the creation processing of *manufacture_plan*; Plan (O) shows the processing to output the query results into csv files for the forms shown in Fig. 2 and Fig. 3; Plan (P) shows the output processing of these forms from the csv files. Moreover, since the SQL statement and mongo shell method need much more descriptions than other instructions, we also show their individual volume in this table.

The function indicated by (1) can be configured only the SQL statements in MySQL. However, in MongoDB, it is necessary to use JavaScript in addition to the mongo shell. In this case, the number of source lines of the latter was about 3.5 times that of the former. The processing indicated by Plan (C) and Plan (O) of (2) cannot be described only with SQL statements in MySQL, so it had to be described with the stored procedures and stored functions. MongoDB was the same as above. In this case, the latter source line number was less than twice that of the former.

The processing indicated by Plan (P) in (2) is the printing of a form, and there is no database access. So, this processing is common to MySQL and MongoDB. That is, the processing to output the form consists of data extraction from the database indicated by Plan (O) and printing as of Plan (P). In this case, the ratio of Plan (P) was 91% in MySQL and 85% in MongoDB.

4.3 Comparative Evaluations of Elapsed Time

We executed each processing mentioned in Section 2.1 on the standalone PC environment to evaluate each elapsed time comparatively. Here, we modified each processing to execute only the database access including the data format operations as the batch file, that is, the following processing are excluded: setting the parameters, printing of forms and so on. So, as for the form output processing, since there was no processing for the database manipulation, we excluded it. The execution environment is as follows. CPU is i7-6700 (3.41GHz); memory is 16GB; disk is SSD memory of 512GB; OS is Windows 10. We adopted MySQL (Ver. 5.7.12), (MongoDB (Ver. 3.4.3) for the database.

We show the evaluation results in Table 3. "Material" shows the elapsed time of the material calculation shown by (a) in Table 2, and it includes the manipulation (a), (c), (d) mentioned in Section 2.2. The elapsed time of MongoDB was about 11 times that of MySQL. "Calendar" and "Plan" are the part of the production planning function: the former creates *calendar*, and as for Mysql, it was executed by stored pro-

cedure with the iterative manipulation shown in (b) of Section 2.2; similarly, the latter creates *manufacture_plan* by SQL statement with joining 3 tables, *calendar*, *order* and *spec*. The elapsed time of MongoDB was about 0.5 and 83 times that of MySQL respectively. “Image (in)” shows the case of insertion of the actual pictures of product shelves into database from the disk: the number of images is 340, and the size of each image is from 0.9 MB to 2.9 MB; “Image (out)” shows the case opposite to the previous case, that is, the same data is queried from the database to store into files. In this processing, MySQL can execute all the data manipulations by connecting once similar to the above-mentioned process. However, since the mongofiles command must be executed for each file as the Windows command individually, the connect operation occurs for each file. In this case, the elapsed time of MongoDB was about 11 and 324 times that of MySQL respectively.

By the way, we separated the fields of the images and videos from the inventory table (*inventory*), and composed the individual table (*image*) even in MySQL. This is due to the results of the preliminary study, in which we gathered all these fields to one table, and found it made the extreme delay. That is, to confirm the inventory, firstly we saved the inventory image shot in order of the shelf ID to this table; then, updated the shelf ID (*shelf_id*) according to the image order. However, this update operation took more than 20 minutes for the above-mentioned 340 data, and it was too long for the operations. Here, it was pointed out that the instances of LONG-BLOB should not in the query results if it was not really necessary [10]. However, as a result of this preliminary study, we found that the extreme latency occurs not only for this pointed out case, but also for the case where the images and video column was not included in the data manipulation.

5 DISCUSSIONS

Next, we discuss on the evaluation results. First, as for the target production management system, we found that all the functions implemented with MySQL can be implemented by using MongoDB. However, some SQL statement had to be described with JavaScript in addition to the mongo shall, that is, the join operation and so on. As a result, as shown in (1) of Table 2, the number of source lines greatly increased compared to the description of SQL statements in MongoDB. Meanwhile, in the case where the program such as stored procedure was used in MySQL, the ratio of increase was small as shown in (2) Plan (C) of Table 2.

On the other hand, as shown in (2) Plan (O) and (P) of Table 2 which is the the form output case, the ratio of description of data manipulation was very small in the actual systems. Here, the other parts include the form creation, entry screen operation and so on. Therefore, from the viewpoint of the overall system development man-hour, we consider that the importance of the selection concerning both will be small.

Incidentally, as shown at the end of section 4.3, in order to maintain the performance, it is necessary to separate the image and video data to the individual table even in MySQL. That is, we found that the same configuration as GridFS of MongoDB is necessary in MySQL, too.

Second, as for the performance that we indicated by the elapsed time, we found there was the extremely large difference between MongoDB and MySQL depending on the type of data manipulation. MongoDB was extremely lower performance than MySQL in the join operation of three tables and the small image query manipulation as shown in Table 3. On the other hand, as shown in Fig. 5, it was extremely faster in the large video data manipulation.

Therefore, as for the target production management system, we currently consider we should substitute MongoDB for only some data manipulation such as the video data manipulation, in which we have the performance problem with MySQL. In the next stage, we will plan to expand the processing, to which we apply MongoDB to improve the system performance, by using the distributed data environment such as the sharding and improving the database structure.

6 CONCLUSIONS

We have the problem in our production management system which is implemented by using MySQL, that is, it is necessary to streamline the inventory management. So, we conceived to use the images and videos to confirm the inventory state from the office. And, in our previous study we had already found that the performance to manipulate the large video is extremely faster in MongoDB than in MySQL.

In this study, we conducted the comparative evaluations between MongoDB and MySQL with the aim of replacing the system with MongoDB. As a results, we found there was the extremely large difference between MongoDB and MySQL depending on the type of data manipulation, and the both have challenges. Therefore, we concluded that we should substitute MongoDB for only some data manipulation such as the large video data manipulation.

For the future challenge, we will expand the application area of MongoDB by using sharding and improving the data structure.

Acknowledgments

This work was supported by JSPS KAKENHI Grant Number 15K00161.

REFERENCES

- [1] R. Arora, and R. R. Aggarwal, “Modeling and querying data in MongoDB,” *International Journal of Scientific and Engineering Research*, Vol. 4, No. 7, pp. 141–144 (2013).
- [2] K. Banker, “MongoDB in Action,” Manning Pubns Co. (2011).
- [3] M. Chen, S. Mao, and Y. Liu, “Big data: A survey,” *Mobile Networks and Applications*, Vol. 19, No. 2, pp. 171–209 (2014).
- [4] C. Győrödi, R. Győrödi, G. Pecherle, and A. Olah, “A comparative study: MongoDB vs. MySQL,” *13th International Conference on EMES*, pp. 1–6 (2015).
- [5] S. Hiremath, G. Yang, and K. Mankodiya, “Wearable Internet of Things: Concept, architectural components

- and promises for person-centered healthcare,” EAI 4th International Conference on Wireless Mobile Communication and Healthcare (2014).
- [6] T. Kudo, M. Ishino, K. Saotome, and N. Kataoka, “A Proposal of Transaction Processing Method for MongoDB,” *Procedia Computer Science*, Vol 96, pp. 801–810 (2016).
- [7] T. Kudo, Y. Ito, and Y. Serizawa, “An Application of MongoDB to Enterprise System Manipulating Enormous Data,” *Proceedings of International Workshop on Informatics (IWIN2016)*, pp. 211–218 (2016).
- [8] K. Nagasawa, and T. Kudo, “Development of Mobile Quotation System Utilizing Tablet and MongoDB,” *Proc. of 2017 IEICE General conference, D-9-23*, p. 113 (2017) (In Japanese).
- [9] Oracle Corp., “Chapter 23 Stored Programs and Views,” <https://dev.mysql.com/doc/refman/5.7/en/stored-programs-views.html> (referred June 6, 2017).
- [10] Oracle Corp., “11.4.3 The BLOB and TEXT Types,” <https://dev.mysql.com/doc/refman/5.7/en/blob.html> (referred June 6, 2017)..
- [11] D.R. Rebecca, and I. E. Shanthi, “A NoSQL Solution to efficient storage and retrieval of Medical Images,” *International Journal of Scientific & Engineering Research*, Vol. 7, No. 2, pp. 545–549 (2016).
- [12] E. Redmond, and J.R. Wilson, “Seven Databases in Seven Weeks: A guide to Modern Databases and the NoSQL Movement,” *Pragmatic Bookshelf* (2012).
- [13] K. Seguin, “The Little MongoDB Book” (2011), <http://openmymind.net/mongodb.pdf> (referred May 5, 2017).
- [14] S. P. Singh, “Production and Operation Management,” *Vikas Publishing House Pvt Ltd* (2014).