

# IWIN2021



International Workshop on Informatics

Proceedings of  
International Workshop on Informatics

September 12-13, 2021  
Virtually Obama, Fukui



Sponsored by Informatics Society





# IWIN2021



International Workshop on Informatics

Proceedings of  
International Workshop on Informatics

September 12-13, 2021  
Virtually Obama, Fukui



Sponsored by Informatics Society

Publication office:

Informatics Laboratory

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

Publisher:

Tadanori Mizuno, President of Informatics

ISBN:

Society 978-4-902523-48-5

Printed in Japan

# Table of Contents

## **Session 1: AI I**

**( Chair: Katsuhiko Kaji ) ( 9:20 - 10:35, Sep. 12 )**

- (1) The proposal on the positioning mechanism using surveillance cameras ..... 3  
Shin-ichi Yamamoto, Ryuji Kubota, Tetsuya Yokotani, Yuichi Tokunaga,  
Masashi Saito, Hironao Kawamura, Hayato Kanayama
- (2) Contour Generation for Object Detection Utilizing Cycle-GAN with Error Monitoring  
..... 13  
Tsukasa Kudo
- (3) Verifications of Influence by Unknown Longer Titles of Work on Robustness of Deep  
Learning NER ..... 23  
Yukihisa Yonemochi, Michiko Oba

## **Session 2: AI II**

**( Chair: Hiroshi Mineno ) ( 10:45 - 12:00, Sep. 12 )**

- (4) Predicting Microclimate Based on Difference from Meteorological Observatory ..... 31  
Genki Nishikawa, Takuya Yoshihiro
- (5) Estimating the best time to see cherry blossoms using SNS and time-series forecasting  
of tweet numbers using machine learning ..... 37  
Tomonari Horikawa, Munenori Takahashi, Masaki Endo, Shigeyoshi Ohno,  
Masaharu Hirota, Hiroshi Ishikawa
- (6) Maintaining Soundness of Social Network by Understanding Fake News  
Dissemination and People's Belief ..... 45  
Risa Kusano, Kento Yoshikawa, Hiroyuki Sato, Masatsugu Ichino,  
Hiroshi Yoshiura

## **Session 3: IoT and ITS**

**( Chair: Yuichi Tokunaga ) ( 13:00 - 14:40, Sep. 12 )**

- (7) Multi-channel Communication of LoRa using Time Division Multiple Access ..... 75  
Sakauchi Ryotaro, Shuto Ishikawa, Hikaru Yabe, Mikiko Sode Tanaka
- (8) Localization Method for an Autonomous Cart as a Guard Robot ..... 59  
Yuya Sawano, Yuto Nagai, Takayuki Suzuki, Ryoza Kiyohara
- (9) A CyReal Approach to Sensor System Development ..... 83  
Kei Hiroi, Akihito Kohiga, Yoichi Shinoda
- (10) Proposal of an efficient downward communication method for a large-scale data  
collection system using MQTT ..... 69  
Fuya Aoki, Koichi Ishibashi, Tetsuya Yokotani

## **Session 4: Network**

**( Chair: Takuya Yoshihiro ) ( 14:50 - 16:30, Sep. 12 )**

- (11) A New Interest Forwarding Method Coping with The Publisher Migration in NDN ...77  
Taichi Iwamoto, Tetsuya Shigeyasu
- (12) A Proposal on mechanisms of ICN with traffic control functions for IoT communication  
..... : 3"  
Atsuko Yokotani, Hiroshi Mineno, Satoshi Ohzahata, Tetsuya Yokotani
- (13) A Study on Assistant Devices for Presentation of Distinctive Viewers' POV  
fin 360-degree Internet Live Broadcasting ..... 89  
Masaya Takada, Yoshia Saito
- (14) An Efficient Large-Scale Video-on-Demand System on Edge Computing Environments  
..... 91"  
Satoru Matsumoto, Tomoki Yoshihisa

## **Session 5: Security and Algorithm**

**( Chair: Tomoya Kitani ) ( 16:40 - 17:55, Sep. 12 )**

- (15) A Design of Plausibly Deniable Distributed File Systems ..... 99  
Ryouga Shibasaki, Hiroshi Inamura, Yoshitaka Nakamura
- (16) A Study on Division Impossibility in the Lightweight N-party Secure Function  
Evaluation for Cloud-Edge Computing Applications ..... 107  
Yutaro Taki, Shigeru Fujita, Norio Shiratori
- (17) Verification of Shell Script Behavior by Comparing Execution Log ..... 111  
Hitoshi Kiryu, Shinpei Ogata, Kozo Okano

## **Session 6: Agricultural IT**

**( Chair: Kei Hiroi ) ( 9:00 - 10:40, Sep. 13 )**

- (18) Proposal of a small robot for agricultural observation ..... 119  
Kenji Terada, Masaki Endo, Takuo Kikuchi, Shigeyoshi Ohno
- (19) Implement of low cost based IoT system in the paddy field to labor-saving and feasible  
study in Otari village, Japan ..... 125  
Kazuma Nishigaki, Kanae Matsui
- (20) A Feature Generation Method for Plant Growth Prediction Using Random Forests ..... 131  
Yosuke Asada, Hiromi Hanada, Takuya Yoshihiro
- (21) Remote monitoring system for mushrooms using LoRa communication ..... 135  
Hikaru Yabe, Mikiko Sode Tanaka

## **Keynote Speech**

**( Chair: Shinji Kitagami ) ( 10:50 - 12:00, Sep. 13 )**

- (I) Digital innovation towards sustainable society ..... 141  
Dr. Akio Yamada, Senior Vice President and Head of NEC Laboratories

## **Session 7: Application I**

**( Chair: Tomoo Inoue ) ( 13:00 - 14:40, Sep. 13 )**

- (22) A System To Directly Feed Back the Audience's Attention Ratio To the Presentation  
Venue ..... 173  
Yuichi Takyo, Katsuhiko Kaji
- (23) Software Edutainment Systems and Analysis of Learners' Data based Docker and  
Edutainment ..... 179  
Ryosuke Tsutsumi, Wei JiuJun, Shinpei Ogata, Masaaki Niimura, Kozo Okano
- (24) Development of a Rainwater Utilization System using LoRaWAN ..... 189  
Shinji Kitagami, Toshihiro Kasai
- (25) Design-thinking information system development methods for the information-impaired  
people ..... 193  
Koji Yamada, Toru Takahashi, Sakiko Kasuya, Nobuhiro Kataoka

## **Session 8: Application II**

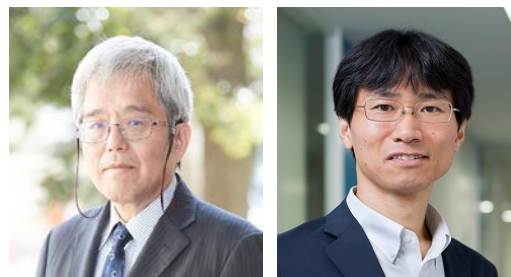
**( Chair: Yoshia Saito ) ( 14:50 - 16:30, Sep. 13 )**

- (26) A proposal of method for collecting daily physical and mental state data using  
communication tools among parenting generation ..... 203  
Kanae Matsui, Kazuma Nishigaki
  
- (27) Development of a Safety Training System for a Portable Grinder that Combines Virtual  
Reality with an Actual Tool ..... 211  
Tomoo Inoue, Asuki Nakanishi
  
- (28) Development of a benchmark system on power-related control by BACnet/IP..... 215  
Kohei Miyazawa, Tetsuya Yokotani, Hiroaki Mukai
  
- (29) Process Improvement of Quantitative Progress Management Process ..... 221  
Akihiro Hayashi





## Message from the General Chairs



It is our great pleasure to welcome all of you to Obama, Fukui, Japan (virtually because of the COVID-19 pandemic), for the 15th International Workshop on Informatics (IWIN 2021). 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, Riga in Latvia, Zagreb in Croatia, Salzburg in Austria, and Hamburg in Germany, Wakayama in Japan (virtually) respectively.

In IWIN 2021, 29 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 Artificial Intelligence, Internet of Things (IoT) and Intelligent Transport System (ITS), Network, Security and Algorithm, Agricultural IT, and Application. IWIN2021 will also welcome one keynote speaker: Dr. Akio Yamada, Senior Vice President and Head of NEC Laboratories, NEC Corporation. We really appreciate his 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 2021.

September 2021

Shinji Kitagami  
Tomoki Yoshihisa

# Organizing Committee

## General Co-Chairs

Shinji Kitagami (Fukui University of Technology, Japan)

Tomoki Yoshihisa (Osaka University, Japan)

## Steering Committee

Hitoshi Aida (Tokyo University, Japan)

Toru Hasegawa (Osaka University, Japan)

Teruo Higashino (Kyoto Tachibana University, Japan)

Tadanori Mizuno (Aichi Institute of Technology, Japan)

Jun Munemori (The Open University of Japan, 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

Katsuhiro Kaji (Aichi Institute of Technology, Japan)

## Financial Chair

Tomoya Kitani (Shizuoka University, Japan)

## Publicity Chair

Yoshitaka Nakamura (Kyoto Tachibana University, Japan)

## Program Committee

Akihiro Hayashi

(Shizuoka Institute of Science and Technology, Japan)

Akihito Hiromori (Osaka University, Japan)

Akira Uchiyama (Osaka University, Japan)

Fumiaki Sato (Toho University, Japan)

Hideki Goromaru (Chiba Institute of Technology, Japan)

Hideyuki Takahashi (Tohoku Gakuin University, Japan)

Hironobu Abe (Tokyo Denki University, Japan)

Hiroshi Inamura (Future University Hakodate, Japan)

Hiroshi Mineno (Shizuoka University, Japan)

Hiroshi Sugimura (Kanagawa Institute of Technology, Japan)

Hiroshi Yoshiura (Kyoto Tachibana University, Japan) Kanae

Matsui (Tokyo Denki University, Japan)

Katsuhiro Naito (Aichi Institute of Technology, Japan)

Kazuaki Nimura (Fujitsu, Japan)

Kazuyuki Iso (NTT, Japan)

Kei Hiroi (Kyoto University, Japan)

Keiichi Abe (Kanagawa Institute of Technology, Japan)

Kozo Okano (Shinshu University, Japan)

Makoto Imamura (Tokai University, Japan)

Masaaki Shirase (Future University Hakodate, Japan)

Masaji Katagiri (iU, Japan)

Masakatsu Nishigaki (Shizuoka University, Japan)

Masaki Endo (Polytechnic University, Japan)

Masaki Nagata (Shizuoka University, Japan)

Masashi Saito (Kanazawa Institute of Technology, Japan)

Michiko Oba (Future University Hakodate, Japan)

Mikiko Sode Tanaka	Tetsuya Yokotani (Kanazawa Institute of Technology, Japan)
(Kanazawa Institute of Technology, Japan)	Tomoo Inoue (University of Tsukuba, Japan)
Minoru Kobayashi (Meiji University, Japan)	Tomoya Kitani (Shizuoka University, Japan)
Naoya Chujo (Aichi Institute of Technology, Japan)	Tsukasa Kudo
Ryozo Kiyohara (Kanagawa Institute of Technology, Japan)	(Shizuoka Institute of Science and Technology, Japan)
Satoru Matsumoto (Osaka University, Japan)	Yoh Shiraishi (Future University Hakodate, Japan)
Shigemi Ishida (Future University Hakodate, Japan)	Yoshia Saito (Iwate Prefectural University, Japan)
Shigeyoshi Ohno (Polytechnic University, Japan)	Yoshiaki Terashima (Soka University, Japan)
Shinichiro Mori (Chiba Institute of Technology, Japan)	Yoshinobu Kawabe (Aichi Institute of Technology, Japan)
Takaaki Umedu (Shiga University, Japan)	Yoshitaka Nakamura (Kyoto Tachibana University, Japan) Yu
Takaya Yuizono	Enokibori (Nagoya University, Japan)
(Japan Advanced Institute of Science and Technology)	Yuichi Bannai (Kanagawa Institute of Technology, Japan)
Takayasu Yamaguchi (Akita Prefectural University, Japan)	Yuichi Tokunaga (Kanazawa Institute of Technology, Japan)
Takuya Yoshihiro (Wakayama University, Japan)	Yuki Koizumi (Osaka University, Japan)
Tetsushi Ohki (Shizuoka University, Japan)	Yusuke Gotoh (Okayama University, Japan)
Tetsuya Shigeyasu	Yusuke Ichikawa (NTT Laboratory, Japan)
(Prefectural University of Hiroshima, Japan)	Yusuke Takatori (Kanagawa Institute of Technology, Japan)



Session 1:

AI I

( Chair: Katsuhiko Kaji )



# The proposal on the positioning mechanism using surveillance cameras

Shin-ichi Yamamoto<sup>\*</sup>, Ryuji Kubota<sup>\*\*</sup>, Tetsuya Yokotani<sup>\*\*</sup>, Yuichi Tokunaga<sup>\*\*\*</sup>, Masashi Saito<sup>\*\*\*</sup>,  
Hironao Kawamura<sup>\*\*\*\*</sup> and Hayato Kanayama<sup>\*\*\*\*</sup>

<sup>\*</sup> Graduate School of Electrical Engineering and Electronics,  
Kanazawa Institute of Technology, Japan  
c6000590@planet.kanazawa-it.ac.jp

<sup>\*\*</sup> Department of Electrical Engineering and Electronics, College of Engineering  
Kanazawa Institute of Technology, Japan  
b1801701@planet.kanazawa-it.ac.jp  
yokotani@neptune.kanazawa-it.ac.jp

<sup>\*\*\*</sup> Department of Management Systems,  
College of Informatics and Human Communication  
Kanazawa Institute of Technology, Japan  
{y.tokunaga, msaito}@neptune.kanazawa-it.ac.jp

<sup>\*\*\*\*</sup> Engineering Research & Development Center,  
Hokuriku Electric Power Company, Japan  
{h.kawamura, kanayama.hayato}@rikuden.co.jp

**Abstract** – We propose a positioning mechanism applying object recognition and pixel coordinates using surveillance cameras. This mechanism can be applied in various fields, such as obstacle detection systems on roads. The following steps are invoked in the proposed mechanism. Step 1 is calculating the distance per pixel coordinate using the two reference objects. Step 2 is conducting object recognition for positioning the target object to derive the image pixel coordinates. Step 3 is converting the image pixel coordinates into real pixel coordinates using the camera angle and height. Step 4 is calculating the target object position according to the real pixel coordinates of the target object on the road. In this paper, we also describe a prototype system that includes the proposed mechanism. We then report a field trial using electrical poles with surveillance cameras to confirm the feasibility of the proposed mechanism.

**Keywords:** positioning, AI, Image processing, trilateration, GPS, ToA, RSSI, WiFi Positioning

## 1 INTRODUCTION

Technology for the positioning of humans and objects is used in many services, and various positioning mechanisms are being researched worldwide [1][2]. It is important for infrastructure to be easily constructed in the positioning mechanism. Furthermore, it is important that the functions be easily customizable for application to various services [3].

In this study, we propose a positioning mechanism that uses surveillance cameras as a mechanism that can easily construct the infrastructure.

With this mechanism, object recognition using AI is applied to object detection, and pixel coordinates are used to position an object.

In Section 2, we investigate the time of arrival (ToA) and received signal strength indicator (RSSI), which are typical positioning mechanisms using wireless communication. In Section 3, the superiority of the proposed method in comparison to the positioning mechanisms investigated in Section 2 is described.

In addition, in this study, we prototyped a system to which the proposed method was applied and evaluated its practicality in the field.

## 2 RELATED WORK

Currently, many of the positioning mechanisms used are mechanisms that use wireless communication [4].

For a comparison with the proposed method, we investigated location identification using the time of arrival (ToA) and receiver signal strength indicator (RSSI), which are typical methods [5][6].

### 2.1 Time of Arrival (ToA)

The time of arrival (ToA) is a method that uses the signal arrival time [7]. GPS positioning is a typical example of using the ToA. In the case shown in Figure 3, assuming that the transmission time of each transmission wave of  $A$  anchors is  $t_k$  ( $k = 1, 2, \dots, A$ ), the reception time at the target is  $t$ , the radio wave propagation speed is  $v$ , and the target starts from the  $k$ th anchor. Distance  $d_k$  is written as

$$\begin{aligned} d_k &= v(t - t_k) \\ &= 3 \times 10^8 (t - t_k) \end{aligned} \quad (2.2.1)$$

Assume that a target  $(x, y)$  whose position is unknown and an anchor  $(x_k, y_k)$ ,  $k = 1, 2, \dots, A$  whose position is

known to **A** exist within a two-dimensional space. The distance  $d_k$  from the  $k$ th anchor to the target is

$$d_k = \sqrt{(x - x_k)^2 + (y - y_k)^2} + S \quad (2.1.2)$$

Here, **S** is the effect of the time error between the transmitter and receiver. The clock times of the receiver and transmitter do not always match. Therefore, the accuracy was improved by calculating **S** to correct the influence of the time error. As shown in Figure 3, when **A** = 3, the equations calculating **x**, **y**, **S** from (2.1.1) and (2.1.2) are solved to find the position [6][8].

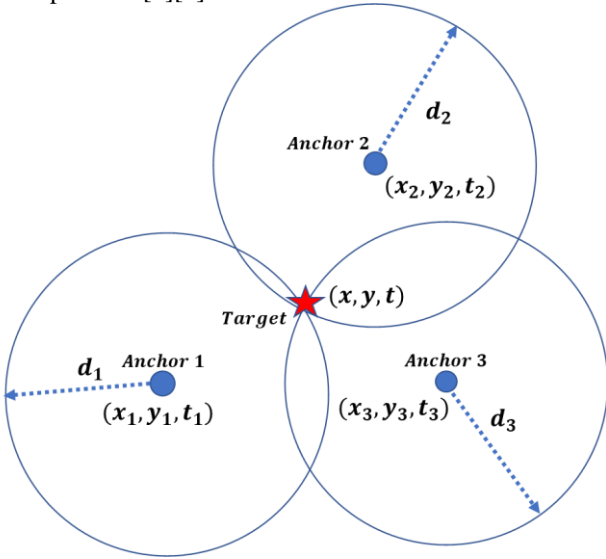


Figure 2 ToA positioning mechanism

## 2.2 Received Signal Strength Indicator (RSSI)

The Received Signal Strength Indicator (RSSI) is a positioning method that uses the signal strength attenuation. WiFi positioning is a typical example of using the RSSI [7], which is a logarithmic representation of the magnitude of the signal strength of a signal sent from a sender node when it is received by a receiver node, mainly during wireless communications such as wireless LAN and Bluetooth, using 1 mW as a reference. The unit of the RSSI is in decibels per milliwatt (dBm), and in most cases, the received power is less than 1 mW, and thus the RSSI is negative. When radio waves are emitted equally in all directions from the antenna of a wireless device, it is known that the power density of a radio wave at a distance **D** [m] is inversely proportional to the surface area of a sphere with **D** as its radius. Therefore, when the transmitting power of a radio wave is **P** [mW], the radio wave strength **P<sub>D</sub>** [mW/m<sup>2</sup>] at distance **D** [m] can be expressed as in (2.2.1) [9].

$$P_D = \frac{P}{4\pi D^2} \quad (2.2.1)$$

**RSSI** expresses the magnitude of the received power **P<sub>D</sub>** at the logarithmic scale with 1 mW as a reference.

Equation (2.2.1) in RSSI format becomes (2.2.2) [9].

$$\begin{aligned} RSSI &= 10 \log_{10} \left( \frac{P}{4\pi D^2} \right) \\ &= 10 \log_{10} \left( \frac{P}{4\pi} \right) - 10 \log_{10}(D^2) \end{aligned} \quad (2.2.2)$$

When the **RSSI** value at a distance of 1 m from the transmitter is set to **RSSI<sub>0</sub>**, (2.2.3) is calculated as follows [9]:

$$\begin{aligned} RSSI &= RSSI_0 - 10 \log_{10}(D^2) \\ &= RSSI_0 - 20 \log_{10}(D) \end{aligned} \quad (2.2.3)$$

As shown in (2.2.3), the RSSI value is negatively correlated with the distance because the radio wave sent out from a node attenuates with distance. Therefore, the distance can be estimated by measuring the RSSI values. The constant value of 20 on the right-hand side of (2.2.3) is a theoretical value, and it is known that the actual value varies depending on the location; thus, it is set as the RSSI attenuation constant **N**, and (2.2.4) is calculated by transforming (2.2.3) [9].

$$D = 10^{-\frac{RSSI - RSSI_0}{N}} \quad (2.2.4)$$

where

- D** = distance [m];
- RSSI** = Measured RSSI value [dbm];
- RSSI<sub>0</sub>** = Unit RSSI value [dbm];
- N** = RSSI damping constant (Theoretically 20 ).

When calculating the distance information from the RSSI, the calculation is applied according to (2.2.4). The RSSI attenuation constant **N** should be an appropriate value for each location.

When the radio signals of three anchors whose positions are known can be received, it is possible to calculate the distance based on the RSSI and estimate the position using trilateration and the ToA. However, when there are obstacles between nodes, the radio waves are attenuated, causing errors in the RSSI values to be measured, which can easily lead to errors in the distance and estimated position [10].

Methods used to compensate for this radio attenuation have also been investigated [9][11].

## 3 THE PROPOSED MECHANISM

The mechanism for calculating the distance through wireless communication and using trilateration investigated in Section 2 requires preparing places where at least three transmitters can be installed and installing the receiver on the object to be positioned.

However, the proposed positioning mechanism using surveillance cameras can complete the positioning with only one camera, and because it is not necessary to add receivers



to the object to be positioned, the infrastructure can be easily constructed.

With the proposed method, positioning is conducted using two reference objects whose pixel coordinates and positions are known along the XY plane with the origin directly below the camera. Figure 3 shows the positioning of the proposed mechanism.

In the positioning mechanism, after positioning in the diagonal direction ( $Y = X$ ) using the pixel coordinates in the vertical direction, positioning in the plane is performed using the pixel coordinates in the horizontal direction. Figure 4 shows the positioning of this mechanism using images.

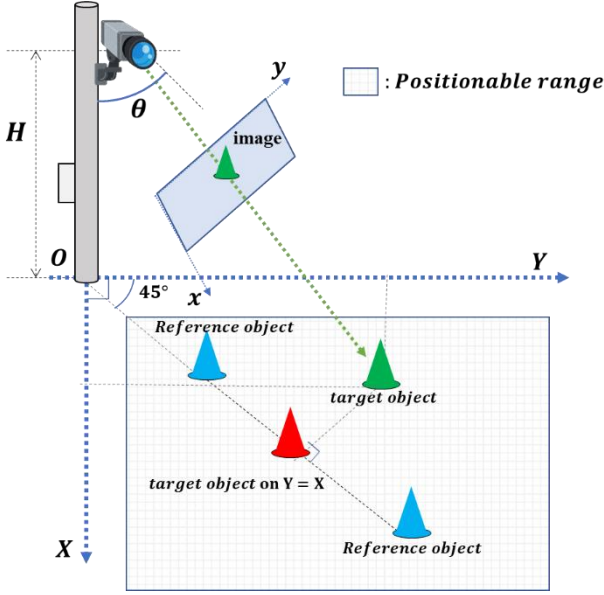


Figure3 Positioning in the proposed mechanism.

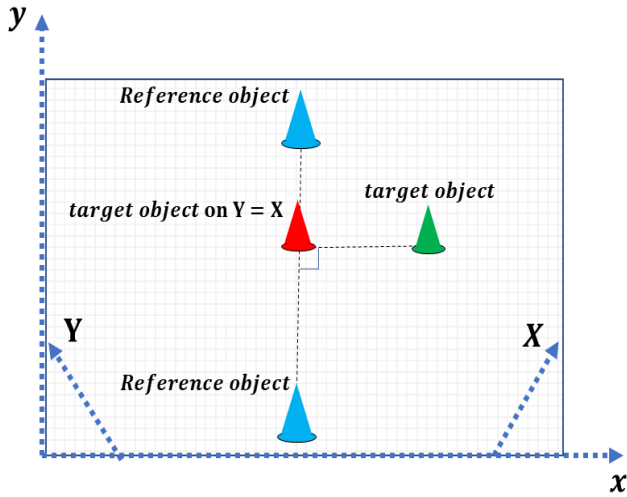


Figure4 The positioning of this mechanism using images

#### 4 DERIVATION OF CONVERSION FORMULA BETWEEN THE IMAGE POSITION AND THE REAL POSITION

To apply one-dimensional positioning, the derivation of the conversion formula between the image position and the real position is described. Figure 5 shows the image position and the real position when viewed from the horizontal direction. The known parameters are the camera height  $HO$  and angle  $\angle KHO = \theta$ .

Let  $A$  be the real position of the object to be positioned. Further, let  $A'$  be the image position of the object to be positioned. Furthermore, we let  $\angle HAO = \alpha$ .

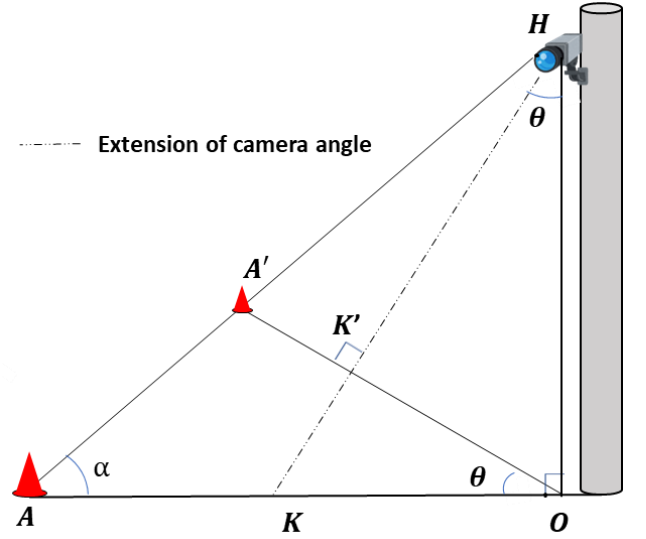


Figure 5 The image position and the real position when viewed from the horizontal direction

From the similarity of  $\triangle KOK'$  and  $\triangle KHO$ ,

$$\angle KOK' = \angle KHO = \theta. \quad (4.1)$$

From the exterior angle theorem,

$$\angle HA'O = \theta + \alpha. \quad (4.2)$$

From  $\angle AOH = 90^\circ$  in  $\triangle AHO$ ,

$$\angle A'HO = 90^\circ - \alpha. \quad (4.3)$$

From the law of sines in  $\triangle A'HO$ ,

$$\frac{A'O}{\sin \angle A'HO} = \frac{HO}{\sin \angle HA'O} \quad (4.4)$$

$$\frac{A'O}{\sin(90^\circ - \alpha)} = \frac{HO}{\sin(\alpha + \theta)} \quad (4.5)$$

$$\frac{A'O}{\cos \alpha} = \frac{HO}{\sin \alpha \cos \theta + \cos \alpha \sin \theta} \quad (4.6)$$

$$\frac{A'O}{\cos \alpha} = \frac{HO}{\cos \alpha \left( \frac{\sin \alpha}{\cos \alpha} \cos \theta + \sin \theta \right)} \quad (4.7)$$

$$A'O = \frac{HO}{\tan \alpha \cos \theta + \sin \theta} \quad (4.8)$$

$$\tan \alpha = \frac{\frac{HO}{A'O} - \sin \theta}{\cos \theta} \quad (4.9)$$

If  $\tan \alpha$  is found in  $\triangle AHO$ ,

$$\tan \alpha = \frac{HO}{AO} \quad (4.10)$$

From (4.9) and (4.10),

$$\frac{\frac{HO}{A'O} - \sin \theta}{\cos \theta} = \frac{HO}{AO} \quad (4.11)$$

Therefore, the following two conversion formulas are derived.

These two conversion formulas are represented by  $f_1(\mathbf{X})$  and  $f_2(\mathbf{X})$ :

$$f_1(AO) = A'O = \frac{1}{\frac{\cos \theta}{AO} + \frac{\sin \theta}{HO}} = \frac{\sec \theta}{\frac{1}{AO} + \frac{\tan \theta}{HO}} \quad (4.12)$$

$$f_2(A'O) = AO = \frac{\cos \theta}{\frac{1}{A'O} - \frac{\sin \theta}{HO}} = \frac{1}{\frac{\sec \theta}{A'O} - \frac{\tan \theta}{HO}} \quad (4.13)$$

Therefore, the conversion formula between the image position and the real position can be derived.

## 5 THE POSITIONING MECHANISM

Let the reference points  $A_1$  and  $A_2$  have the exact position and coordinates on the image in advance.

In the following, the uppercase letters are distances and the lowercase letters are the pixel coordinates.

That is,  $A_1O, A_2O, a'_1o', a'_2o'$  are known. In addition, let  $T$  be the vertical position of the object to be positioned.

Figure 6 shows the positioning viewed from the horizontal direction, including the two reference points.

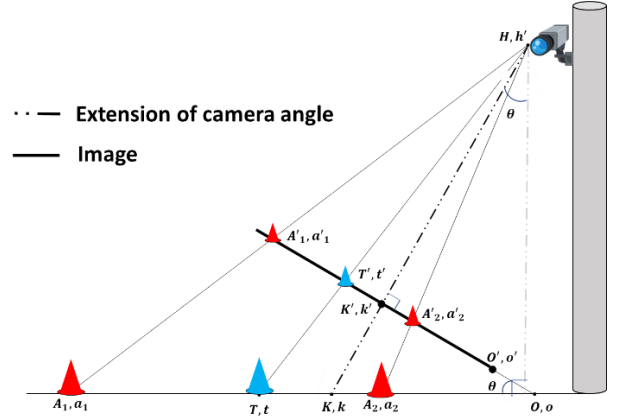


Figure 6 The positioning viewed from the horizontal direction including the two reference points.

First, the distance per unit pixel is calculated.

Using (4.12), the distance between  $A'_1$  and  $A'_2$  is

$$A'_1A'_2 = f_1(A_1O) - f_1(A_2O) = A'_1O - A'_2O. \quad (5.1)$$

The distance between  $a'_1$  and  $a'_2$  is

$$a'_1a'_2 = a'_1o' - a'_2o'. \quad (5.2)$$

Using (5.1) and (5.2), the distance per unit pixel coordinate  $m_p$  is calculated as follows:

$$m_p = \frac{A'_1A'_2}{a'_1a'_2} [m/pixels]. \quad (5.3)$$

From the pixel coordinate per unit distance,  $p_m$  is the reciprocal of (5.3):

$$p_m = \frac{1}{m_p} [pixels/m]. \quad (5.4)$$

Calculate  $O'O$  in Figure 5. Here,  $O'O$  is calculated by the difference between  $K'O$  and  $K'O'$ . When  $K'O'$  is calculated,

$$K'O = HO \sin \theta. \quad (5.5)$$

After calculating  $K'A'_2$ ,  $k'a'_2$ , and  $k'o'$  in order, calculate  $K'O'$ . Here,  $K'A'_2$  is calculated by the difference between  $K'O$  and  $A'_2O$ :

$$K'A'_2 = K'O - A'_2O. \quad (5.6)$$

Here,  $k'a'_2$  is calculated using  $K'A'_2$  and  $p_m$ :

$$k'a'_2 = K'A'_2 \times p_m. \quad (5.7)$$

Here,  $k'o'$  is calculated using  $k'a'_2$  and  $a'_2o'$ .

$$k'o' = k'a'_2 + a'_2o'. \quad (5.8)$$

Here,  $\mathbf{K}'\mathbf{O}'$  is calculated using  $\mathbf{k}'\mathbf{o}'$  and  $\mathbf{m}_p$ :

$$K'O' = k'o' \times m_p. \quad (5.8)$$

From (5.5) and (5.8)

$$O'O = K'O - K'O'. \quad (5.9)$$

Calculate  $\mathbf{T}\mathbf{O}$  by applying (4.13) to  $\mathbf{T}'\mathbf{O}$ .

Here,  $\mathbf{T}'\mathbf{O}$  is calculated using  $\mathbf{t}'\mathbf{o}'$  (Y coordinate of the object to be positioned) and  $\mathbf{m}_p$ , and  $\mathbf{T}\mathbf{O}$  is calculated from the sum of  $\mathbf{T}'\mathbf{O}'$  and  $\mathbf{O}'\mathbf{O}$ .

$$\mathbf{T}'\mathbf{O}' = \mathbf{t}'\mathbf{o}' \times \mathbf{m}_p \quad (5.10)$$

$$T'O = T'O' + O'O \quad (5.11)$$

$$T\mathbf{O} = f_2(T'\mathbf{O}) \quad (5.12)$$

Next, we will consider the positioning on a real plane. Positioning on a plane is shown in Figure 7.

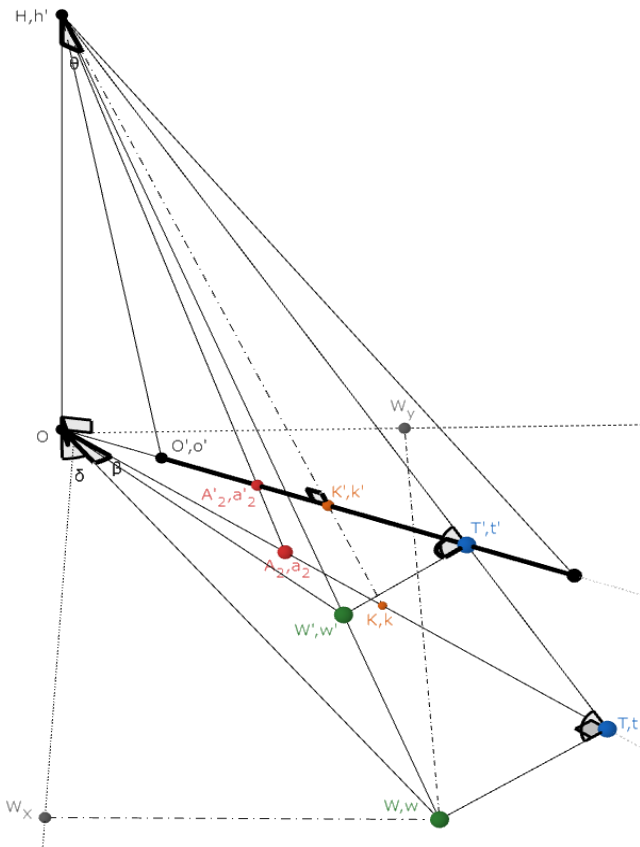


Figure 7 Positioning on a plane

Positioning is conducted on a plane with O as the origin.

Use the similarity between  $\triangle THW$  and  $\triangle T'HW'$  to calculate  $\mathbf{wt}$  from  $\mathbf{w}'\mathbf{t}'$ , and use  $\mathbf{wt}$  and  $\mathbf{to}$  to calculate  $\angle TOW$ .

First,  $\mathbf{w}'\mathbf{t}'$  is calculated based on the difference between the X coordinate  $\mathbf{w}'_x$  of the positioning target object on the image and the X coordinate  $\mathbf{t}_x$  of the reference point.

$$\mathbf{w}'\mathbf{t}' = \mathbf{w}'_x - \mathbf{t}_x \quad (5.13)$$

The values of  $\mathbf{TH}$  and  $\mathbf{T'H}$  are calculated such that  $\Delta \mathbf{THW}$  and  $\Delta \mathbf{T'HW'}$  use the similarity relationship.

In addition,  $\angle TOH = 90^\circ$ , and thus if  $TH$  is calculated using  $HO$  and  $TO$ ,

$$TH = \sqrt{HO^2 + (TO)^2}. \quad (5.14)$$

The value of  $\mathbf{T}'\mathbf{H}$  is calculated using the fact that  $\angle \mathbf{T}'\mathbf{K}'\mathbf{H} = 90^\circ$ , from which  $\mathbf{H}\mathbf{K}'$  and  $\mathbf{T}'\mathbf{K}'$  are also calculated.

$$HK' = HO \cos \theta \quad (5.15)$$

$$\mathbf{T}'\mathbf{K}' = |\mathbf{T}'\mathbf{O} - \mathbf{K}'\mathbf{O}| \quad (5.16)$$

From (5.15) and (5.16),

$$\mathbf{T}'\mathbf{H} = \sqrt{(\mathbf{H}\mathbf{K}')^2 + (\mathbf{T}'\mathbf{K}')^2}. \quad (5.17)$$

From  $\angle OTW = 90^\circ$ , calculate  $\angle TOW$  using **wt** and **to**.

$$wt = w't' \times \frac{TH}{T'H} \quad (5.18)$$

$$to = T\mathbf{O} \times p_m \quad (5.19)$$

$$\beta = \angle TOW = \tan^{-1} \frac{wt}{to} \quad (5.20)$$

Here,  $\mathbf{W}\mathbf{O}$  is calculated using  $\mathbf{T}\mathbf{O}$ .

$$W O = T O \sec \beta \quad (5.21)$$

Here,  $\delta = \angle \mathbf{TO} \mathbf{W}_y$  calculates the angle according to the magnitude of  $\mathbf{t}_x$  and  $\mathbf{w}_x$ .

**IF**  $t_x < w_x$

$$\delta = 45^\circ - \beta$$

**IF**  $t_x > w_x$

$$\delta = 45^\circ + \beta \quad (5.22)$$

When **WO** is indicated by X and Y,

$$W\mathbf{O}_x = W\mathbf{O} \cos \delta$$

$$W O_y = W O \sin \delta \quad (5.23)$$

Therefore, positioning can be performed in the XY coordinate system ( $\mathbf{W}\mathbf{O}_x, \mathbf{W}\mathbf{O}_y$ ) and polar coordinate system ( $\mathbf{W}\mathbf{O}, \delta$ ) with  $\mathbf{O}$  as the origin.

## 6 HOW TO DETERMINE THE HEIGHT AND ANGLE OF THE CAMERA

The angle and height of the camera are important parameters that determine the range and accuracy of the positioning.

In the proposed method, the range in which the positioning is possible is along the XY plane and is the shooting range of the camera. In addition, although the positioning

accuracy is not constant at all positions, the positioning accuracy at each position changes depending on the angle and height of the camera. Therefore, it is necessary to appropriately set the angle and height of the camera when using the proposed method.

This section presents a method for determining the height and angle of the camera to apply the positioning with an appropriate positioning range and the positioning accuracy according to the application.

As the object to be positioned moves farther away, the distance between the image positions narrows. Figure 8 shows the difference between the real position and position on the image.

This change in the spacing of the image affects the positioning accuracy depending on the angle and height of the camera.

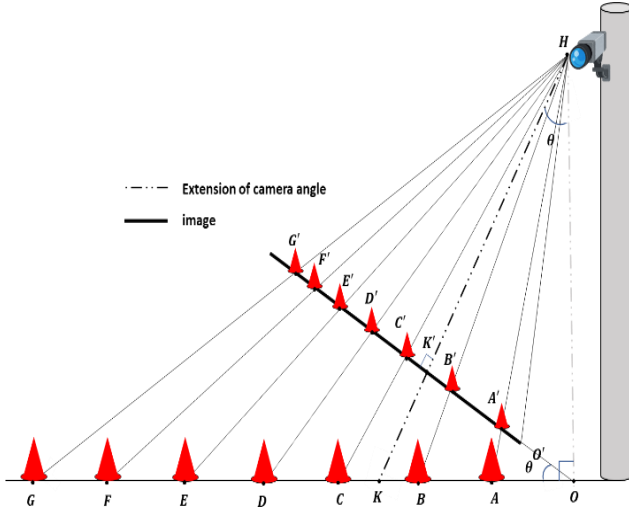


Figure 8 The difference space between the real position and the position on the image

The number of pixels contained in a minute space on the image was evaluated, and the positioning accuracy was derived using a mathematical formula.

From (4.12), if the actual position is  $x$ , it can then be expressed as (6.1).

$$f_1(x) = \frac{\sec \theta}{\frac{1}{x} + \frac{\tan \theta}{HO}} \quad (6.1)$$

Differentiating (6.1), we have

$$\frac{df_1(x)}{dx} = \frac{\sec \theta}{\left(\frac{\tan \theta}{HO}x + 1\right)^2} \quad (6.2)$$

When the size of the image sensor is  $3.6 \times 2.7$  [mm], the focal length  $f$  is 3.6 [mm], and the pixel size is  $640 \times 480$ , the focal length, the size of the image sensor, the distance to the subject, and the range  $L$  where positioning is possible, are similar.

Figure 9 shows the relationship between the image sensor and the shooting range.

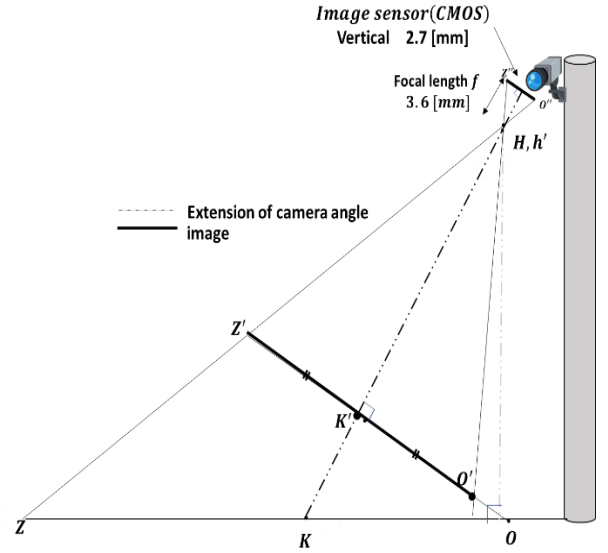


Figure 9 Relationship between image sensor and shooting range

Using this similarity relationship, the distance  $m_p$  per unit pixel coordinate in the image sensor can be calculated.

$$L = \frac{HK' \times 2.7}{f} = \frac{3}{4} \times HO \cos \theta \quad (6.4)$$

$$p_m = \frac{480}{L} = \frac{640}{HO \cos \theta} \quad (6.6)$$

Converting (6.2) into the pixel coordinates,

$$p_{m2} = p_m \times \frac{df_1(x)}{dx} = \frac{640}{HO} \times \frac{\sec^2 \theta}{\left(\frac{\tan \theta}{HO}x + 1\right)^2} \quad (6.7)$$

In addition, the conditions of the range where positioning in the vertical direction is possible are as follows:

$$\begin{aligned} O'O &= K'O - K'O' \\ &= HO \sin \theta - \frac{HO \cos \theta \times 2.7}{\frac{3.6 \times 2}{3 \cos \theta}} \\ &= HO \times \left(\sin \theta - \frac{3 \cos \theta}{8}\right) \end{aligned} \quad (6.8)$$

$$\begin{aligned} Z'O &= K'O + K'O' \\ &= HO \times \left(\sin \theta + \frac{3 \cos \theta}{8}\right) \end{aligned} \quad (6.9)$$

$$\begin{aligned} ZO &= f_2(Z'O) \\ OO &= f_2(O'O) \end{aligned} \quad (6.10)$$

$$OO \leq x \leq ZO \quad (6.11)$$

Using (6.7) and (6.11), the height  $HO$  and angle  $\theta$  of the camera are applied based on the parameters of multiple patterns to evaluate the change in the positioning range and positioning accuracy depending on the position.

Figure 10 shows an example of the relationship between the distance  $x$  and the number of pixels in minute sections owing to the camera height  $HO$  and angle  $\theta$  of multiple patterns. Finally, the graph shows theoretical values.

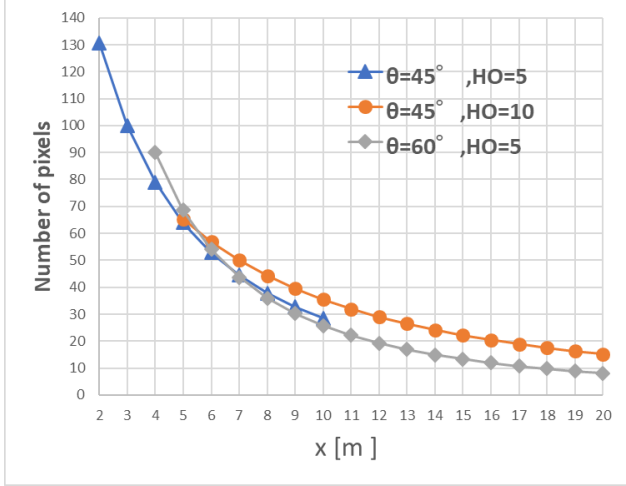


Figure 10 An example of the relationship between the distance  $x$  and the number of pixels in minute sections

As shown in Figure 10, it is necessary to set the height and angle of the camera while considering the trade-off relationship between the positioning range and positioning accuracy. Furthermore, in the same manner, it is necessary to calculate the range  $V$  that enables positioning in the horizontal direction.

$$\frac{HO \cos \theta \times 3.6}{3.6 \times 2} = \frac{HO \cos \theta}{2} \quad (6.12)$$

$$-\frac{HO \cos \theta}{2} \leq V \leq \frac{HO \cos \theta}{2} \quad (6.13)$$

Similarly, in horizontal positioning, the parameters of the height  $HO$  and angle  $\theta$  of the camera affect the positioning range.

In actual positioning, use (6.7), (6.11), and (6.13) to appropriately set the height  $HO$  and camera angle  $\theta$  according to the application. Furthermore, the value of the image sensor is a theoretical value and is not used for the positioning of the proposed method because an error occurs owing to a small individual difference in the image sensor.

## 7 PROTOTYPING AND PERFORMANCE EVALUATION OF THE POSITIONING SYSTEM

A Jetson Nano [12] (Figure 11), which is a development board equipped with a GPU for object recognition, was used to implement the proposed method. Table 1 lists the Jetson Nano specifications [12]. YOLO v3 [13] was used as the

detection method for object recognition. The operating environment of the Jetson Nano is shown in Table 2. A traffic cone (Figure 12) was used as the object to be positioned.

The result of the distance calculation was confirmed on a desktop or tablet PC using a wireless router.

Figure 13 shows the configuration of the prototype system, Table 3 shows the specification of the surveillance cameras used, and Figure 14 shows the state of the object recognition when applying the Jetson Nano.



Figure 11 Jetson Nano[12]

Table1 Jetson Nano specifications

GPU	128-core Maxwell
CPU	Quad-core ARM A57 @ 1.43 GHz
memory	4 GB 64-bit LPDDR4 25.6 GB/s

Table2 The operating environment of Jetson Nano

Jetpak	4.2.2
OS	Ubuntu 18.04
CUDA	10.0.326
cuDNN	7.5.0.56
Python	3.6.9
pillow	8.1.0
numpy	1.19.3
tensorflow-gpu	1.13.1+nv19.4
keras	2.2.4
matplotlib	3.3.3
opencv-python	4.5.1.48

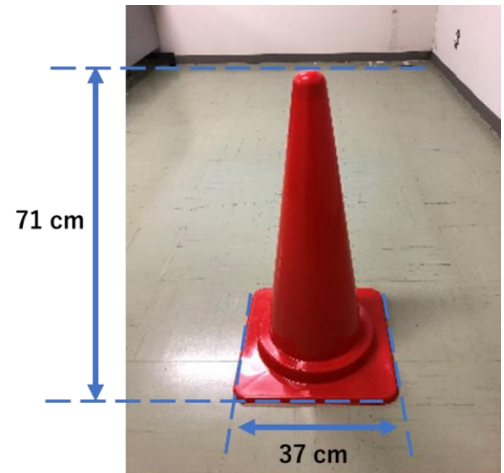


Figure 12 Positioning target object in verification (Road cone)



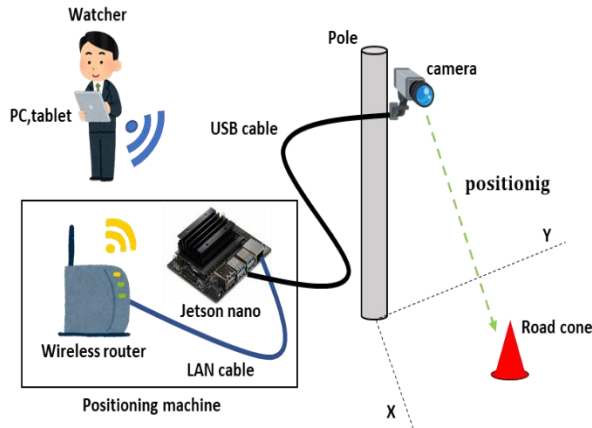


Figure 13 The configuration of the prototype system

Table3 the specification of the surveillance cameras used

Image sensor	1/4" CMOS OV9712
Image sensor size	$3.6 \times 2.7$ mm
lens	3.6mm
Resolution used	640 (H) $\times$ 480 (V) pixels



Figure 14 The state of object recognition in Jetson Nano.

In addition, a performance evaluation of the prototype system used by the proposed mechanism was conducted in the field, as shown in Figure 15.

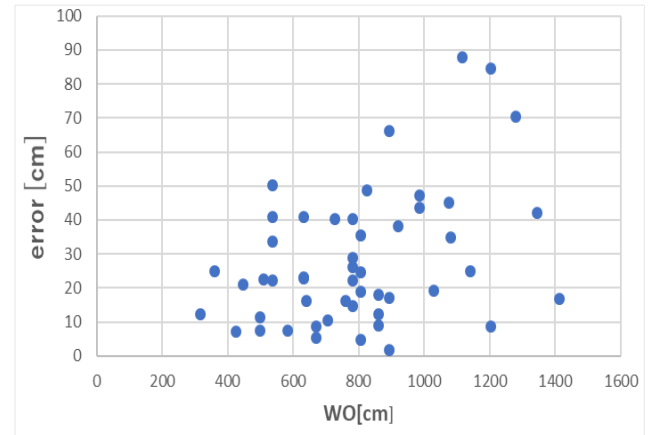
In the evaluation environment, the performance was evaluated at a height  $HO$  of 3.85 m and an angle  $\theta$  of  $56^\circ$ . In this verification, the positioning range is  $10\text{ m} \times 10\text{ m}$ . The accuracy of object recognition is not considered in this verification.



Figure 15 The performance evaluation in the field

In the performance evaluation, as a result of applying the positioning along the XY plane 50 times at random positions within the possible range, an error of 28 cm for  $WO_x$  and 23 cm for  $WO_y$ , occurred on average.

Figure 16 shows the relationship between the distance  $WO$  from the origin  $O$  to the target object and the error. Although no significant correlation between the distance  $WO$  and the error was shown, in some cases, the error increased as the distance increased. In addition, the maximum error was within 1 m.

Figure 16 The relationship between the distance  $WO$  and the error

## 8 CONCLUSION

In this study, we proposed a positioning mechanism using a surveillance camera as a mechanism that can be used to easily construct an infrastructure. With this method, it was found that the height and angle of the camera are important factors related to positioning accuracy. Furthermore, the positioning accuracy can be improved by appropriately defining the height and angle of the camera according to the application.

Furthermore, the object recognition system was operated using a Jetson Nano, the positioning function of the proposed mechanism was added to the system, and the positioning system using the surveillance camera was prototyped.

During the field performance evaluation, the practicality of the proposed mechanism was evaluated using a prototype system, and the error shown was within 30 cm on average along the XY plane, confirming that the proposed mechanism is a practical tool.

With the proposed method, because the positioning mechanism is completed with one surveillance camera, it is easy to introduce the infrastructure, and it is expected that positioning will be possible in poor weather or in places where it is difficult to install radio equipment.

Our future task is to improve the accuracy of the object recognition, which was not considered in this verification. During this verification, object recognition was conducted using a trained model; however, in actual cases, it is also necessary to detect unlearned objects, and thus it is necessary to consider an approach for object detection related to such objects.

## REFERENCES

- [1] L. Mainetti, L. Patrono and I. Sergi, A survey on indoor positioning systems, 2014 22nd International Conference on Software, Telecommunications and Computer Networks (SoftCOM), pp. 111-120 (2014)
- [2] Q. Liu, J. Qiu and Y. Chen, Research and development of indoor positioning, in China Communications, vol. 13, no. Supplement2, pp. 67-79 (2016)
- [3] M. B. Kjærgaard, M.V. Krarup, A. Stisen, T. S. Prentow, H. Blunck, K. Grønbæk, and C.S. Jensen, Indoor positioning using wi-fi—how well is the problem understood?, International Conference on Indoor Positioning and Indoor Navigation. Vol. 28 (2013)
- [4] H. Liu, H. Darabi, P. Banerjee and J. Liu, Survey of Wireless Indoor Positioning Techniques and Systems, in IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), vol. 37, no. 6, pp. 1067-1080 (2007)
- [5] S. Gezici, A survey on wireless position estimation, Wireless personal communications 44.3 (2008)
- [6] O. S. Oguejiofor, A. N. Aniedu, H. C. Ejiofor and A. U. Okolibe, Trilateration based localization algorithm for wireless sensor network, International Journal of Science and Modern Engineering (IJISME) 1.10 (2013)
- [7] S. C. Yeh, W. H. Hsu, M. Y. Su, C. H. Chen, and K. H. Liu, A study on outdoor positioning technology using GPS and WiFi networks, 2009 International Conference on Networking, Sensing and Control pp. 597-601 (2009)
- [8] M. F. Mosleh, M. J. Zaiter and A. H. Hashim, Position Estimation Using Trilateration based on ToA/RSS and AoA Measurement, Journal of Physics: Conference Series. Vol. 1773. No. 1 (2021)
- [9] W. Nakai, Y. Kawahama, and R. Katsuma, Improvement of Positioning Accuracy by Estimating Unit RSSI, Information Processing Society of Japan Kansai Branch Conference Proceedings, 7p (2017)
- [10] A. Bose, and C. H. Foh, A practical path loss model for indoor WiFi positioning enhancement, 2007 6th International Conference on Information, Communications & Signal Processing, pp. 1-5 (2007)
- [11] S. Sadowski, and P. Spachos, Rssi-based indoor localization with the internet of things, IEEE Access 6, vol. 6, pp. 30149-30161 (2018)
- [12] A. Kurniawan, Introduction to NVIDIA Jetson Nano, IoT Projects with NVIDIA Jetson Nano. Apress, Berkeley, CA, pp. 1-6 (2021)
- [13] J. Redmon, and A. Farhadi, Yolov3: An incremental improvement, arXiv preprint arXiv:1804.02767 (2018)





# Contour Generation for Object Detection Utilizing Cycle-GAN with Error Monitoring

Tsukasa Kudo<sup>†</sup>

<sup>†</sup>Faculty of Informatics, Shizuoka Institute of Science and Technology, Japan  
kudo.tsukasa@sist.ac.jp

**Abstract** - In recent years, with the spread of IoT, a huge amount of image data has been input into systems, and it has become necessary to automatically extract the required information from it. So, various studies on object recognition have been carried out, and remarkable development has been achieved especially by utilizing deep learning. Here, when the target area is small in the image, it is necessary to detect the target object and extract its area for its recognition firstly. In this field as well, various studies utilizing image processing and deep learning are being actively conducted, and improvements in efficiency and accuracy have been achieved. However, to specify the target area or contour in a pixel-to-pixel manner, it is necessary to prepare each pair of the original and its ground truth images as training data, which is a practical obstacle. In this study, I propose a method of translating a target image into a contour-enhanced image using Cycle-Consistent Adversarial Networks (Cycle-GAN), which does not require pairs of training images. Furthermore, through experiments, it is shown that each contour of the targets can be detected collectively even for plural dense objects.

**Keywords:** Cycle-GAN, Object detection, Contour detection, Image processing, Image-to-image translation

## 1 INTRODUCTION

At present, a huge amount of data is input as big data from various sensors with the progress of IoT. And, attempts to automatically extract useful information from these data by utilizing deep learning (DL) are widely made, and high discrimination accuracy is achieved [17], [26]. Among such sensors, to monitor targets by images, a large number of cameras such as surveillance, in-vehicle, river, and wearable cameras have been deployed, and necessary information is automatically extracted from various videos and images.

Here, when the target area in the image is relatively small, it is necessary to detect the target object to extract its area, then the target recognition is performed in this area. For example, in face recognition, a face area is extracted from an image using the Haar-Like features, then face recognition is performed in this area to identify the target person [27].

As such a research field, the author has dealt with the theme of automatically extracting the necessary information for inventory management of the parts shelves in machine assemble factories. Since a factory often has thousands or more inventory shelves, efficient data collection and automatic information extraction are required. So, I have been studying a method of having workers wear wearable cameras and automatically collecting images of inventory shelves during their

work. Even with this method, to recognize the target object such as parts, if their areas are small in an image, it is necessary to extract the target area.

For this problem, I focused on the fact that workers pick up parts by hand when they conduct works about inventory operations such as replenishment and shipping. And, using optical flow, which is a representation of the movement of an object between adjacent frames in videos, it was shown that the target object area can be extracted in this case [11]. However, it was a remainder problem to extract the target area for the still target objects, especially from the still image including the plural dense objects, such as the lined-up parts containers in the factory.

Regarding object detection, various methods have been proposed in the conventional image processing field such as area segmentation and contour detection [29]. In addition, in recent years, various methods using DL are proposed [16], [28], and it has been shown that the target area or contour can be extracted in a pixel-by-pixel manner with some methods. However, since these methods need each pair of the original and its ground truth images as training data, it is an obstacle to actual application.

On the other hand, Cycle-Consistent Adversarial Networks (Cycle-GAN) have been proposed, which is a kind of generative adversarial networks (GAN) [32], [5]. And, it has been shown that mutual image translation between two different types of images can be performed by Cycle-GAN such as between photographs of horses and zebras. In addition, just both types of images are necessary for the training data, and they do not need to be paired. This suggests that an image including target objects can be translated into one that emphasizes the contour of the objects using the training data prepared efficiently without making data pairs as above-mentioned, and the target object region can be extracted from it.

In this study, a method to extract contours using a contour extraction model (CE-model) is proposed, which is based on Cycle-GAN. In this method, the following two types of images are translated mutually. One is the original image; another is its contour emphasis image, in which the contour area is represented as so brighter, and other areas are darkened. And, it is shown that the contour of the target object can be obtained from this contour emphasis image generated by the CE-model trained with these images.

In Cycle-GAN, the accuracy of the extracted contour cannot be directly monitored because supervised training using ground truth is not performed. In addition, it has been found that the loss of CycleGAN and the error of the object created by it does not always consistent, through my previous re-

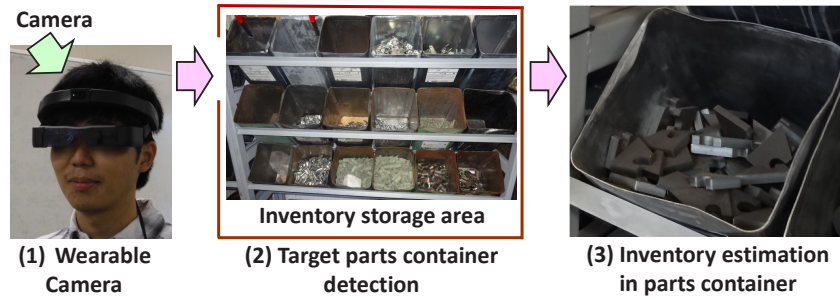


Figure 1: Automatic inventory management using videos

search [12]. So, this method incorporates a function of monitoring this error to stop the Cycle-GAN model training at the optimum timing. Furthermore, through experiments targeting books arranged on bookshelves, it is shown that targets' contours can be extracted collectively from a still image even including plural dense objects by this method.

The remainder of this paper is organized as follows. Section 2 describes related works and motivation for this study, and Sec. 3 proposes a contour extraction method. Section 4 shows the implementation of the experimental system and evaluations, and its results are discussed in Sec. 5. Finally, Sec. 6 concludes this paper.

## 2 RELATED WORKS AND MOTIVATION

Nowadays, with the progress of IoT, it has become possible to connect various cameras to the Internet and easily share and analyze recorded videos. As a result, using cameras such as wearable cameras and in-vehicle cameras, it has become possible to automatically extract necessary information from the scenery we see casually in our daily lives and utilize it efficiently.

Regarding such IoT applications, the author has been working on a study to automate inventory management in machine assembly factories. There are various parts stored in the bulk container shown in Fig. 1 (3), and they cannot be counted visually from the outside. Furthermore, since the number of these containers often reaches several thousand or more, efficient inventory management has been an issue. For example, an attempt to grasp quantities of parts by measuring the weight of each container has been unsuccessful, because the uneven distribution of parts in containers causes a non-negligible error and too many scales were required. For this issue, I focused on the fact that inventory fluctuates only with inventory operations such as the worker's replenishing or shipping of parts, and conceived to extract the necessary inventory information automatically from the videos shot by the wearable camera worn by the worker.

Using smart glasses equipped with the camera shown in Fig. 1 (1), workers can take videos of the target object without extra load and confirm the shot video by displaying it on the see-through glasses. And, I have shown that some necessary information for inventory management could be extracted automatically by applying the multi-class classification and regression model of DL to these videos in my previous study

[9], [10]. By the former, the current position such as entering the inventory storage area shown in Fig. 1 (2) could be determined; by the latter, inventory quantities of parts in the bulk containers shown in (3) could be estimated, even though it could not be counted from the outside visually as above-mentioned.

Furthermore, I have shown even when a target object's area was relatively small in an image, it could be extracted by utilizing the optical flow when the target is moving against the background, such as the worker picking up parts during the work [11]. Supplementally, in this case, the target area has to be extracted first to maintain its recognition accuracy, similar to such as face detection in face recognition. However, extraction of a target area from dense still objects, such as the lined-up bulk containers as shown in Fig. 1 (2) remained an issue. And, this is the subject of this study.

To detect such objects, some methods have been proposed in the conventional image processing fields. In contour tracking, the target's contour is extracted to identify its area; In edge detection, the boundary of the target is specified using a filter or the like; In segmentation using such as mean-shift clustering, pixel value or texture gradient is utilized to determine the target area [29]. But, it has been pointed out that these contour detection methods using image processing become difficult tasks when contours are incomplete or not closed [4].

On the other hand, with the progress of DL in recent years, studies on object detection have been actively conducted, and various methods have been proposed. One is based on the region proposal that detects the bounding boxes where the objects exist, and YOLO (You Only Look Once) achieved high efficiency by detecting images by CNN (Convolutional Neural Network) processing only once; SSD (Single Shot Detector) enabled to detect objects of various sizes with one processing, and RatinaNet improved its efficiency [16], [19], [18], [14], [15]. Furthermore, even for dense objects, some ways are shown to detect the individual object's region, such as detecting new objects repeatedly (IterDet) and separating duplicate regions by post-processing [1], [20]. However, since these region proposals use bounding boxes, namely rectangles, the exact area of targets cannot be specified.

To detect object regions or contours in a pixel-by-pixel manner, various approaches have been studied. A basic method has been proposed in which each pixel is judged whether in an object's contour using CNN [23]. For the methods based

on the region proposal, semantic and instance segmentation have been proposed, which are segmentation for each object's class and object itself respectively [31], [7]. Furthermore, after the method of image-to-image translation with cGANs (conditional GANs) between original and feature images was shown as pix2pix [8], methods using GAN have been studied actively [13], [24], [28], [30]. However, since these methods require pairs of the original and ground truth images as training data, there is an obstacle to their practical use.

Cycle-GAN has been proposed as one of the GANs for image-to-image translation. By using it, images in a domain can be translated to the ones in another domain mutually, and the original paper showed examples of mutual translation between photographs and painter's drawings, summer and winter photographs, and so on [32]. Regarding the above-mentioned obstacle for practical use, Cycle-GAN has an important feature that it is not necessary to prepare each image pair of two domains as training data, that is, it is easier to prepare training data.

In my previous study, I have utilized Cycle-GAN to prepare efficiently the training data for the model to estimate the inventory quantity shown in Fig. 1 (3). Concretely, this model was trained with CG images generated automatically, and inventory was estimated with fake CG images translated from photographs using Cycle-GAN. As a result, it was shown that the estimation accuracy could be improved compared to the case of using the original photographs [12]. However, through this study, it was also found that there was no correlation between the loss of Cycle-GAN in training and the transition of the above-mentioned estimation accuracy, that is, there was an issue when to stop model training.

Similarly, various applications of Cycle-GAN have been proposed. The first is the augmentation of training data, which is used in the fields where it is difficult to generate sufficient training data for DL, such as medical treatment and detection of plant lesions [22], [25]. The second is to translate the original image into an image that is easier to detect objects as a preprocessing, and methods combined with such as YOLO and RetinaNet have been proposed [21], [3]. However, I could not find the study targeting contour detection for still images including dense objects.

The goal of this study is to develop a method to extract the target objects' contours collectively from a still image including plural dense objects such as the lined-up containers' image shown in Fig. 1 (2) using Cycle-GAN. And, its results are assumed to be used for target detection and recognition. For example, by identifying the inventory shelf currently in operation, the target part can be identified. And, its inventory quantity can be estimated using the above-mentioned regression model of my previous study. In addition, the motivation of this study is the question of whether the contours of the target objects can be extracted from contour emphasis images which are translated from original images using Cycle-GAN.

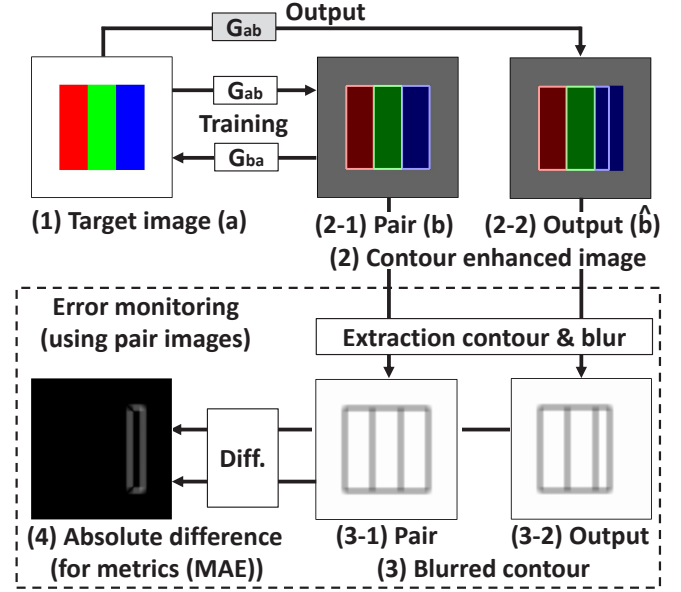


Figure 2: Contour extraction method utilizing Cycle-GAN

### 3 PROPOSAL OF CONTOUR EXTRACTION MODEL

#### 3.1 Contour Generation Applying Cycle-GAN

Figure 2 shows the method of extracting the contours of the target's object from a still image including plural dense objects using Cycle-GAN. Figure 2 (1) shows the image that includes three dense objects. (2-1) is an image that emphasizes the contour of the target's objects in (1). In this study, CE-model utilizing Cycle-GAN is proposed, which is trained with images in these two domains (1) and (2-1). From images in the domain (1), its corresponding contour emphasis images shown in (2-2) are generated using the CE-model.

When the image of Fig. 2 (1) is shown by  $a$  and the image of (2-1) is shown by  $b$ , generators of the CE-model that perform mutual translation between them are shown below.

$$\hat{a} = G_{ba}(b) \quad (1)$$

$$\hat{b} = G_{ab}(a) \quad (2)$$

Here,  $\hat{a}$  and  $\hat{b}$  are fake images of the image  $a$  and  $b$  respectively. For image  $b$ , the CE-model is trained using discriminators that monitor the following three losses same as Cycle-GAN.

$$L_b = \| G_{ba}(b) - a \| \quad (3)$$

$$L_c = \| G_{ab}(G_{ba}(b)) - b \| \quad (4)$$

$$L_i = \| G_{ab}(b) - b \| \quad (5)$$

Here,  $\| \cdot \|$  indicates an error, and similar losses are monitored for  $a$ .  $L_b$  evaluates the error between the fake image  $\hat{a}$  and  $a$ ;  $L_c$  evaluates the reconstruction image generated by applying these two generators sequentially, namely between the fake image of  $b$  and  $b$  itself;  $L_i$  evaluates the identity image, which is generated by the generator for this image itself. In training, these losses are added with a specified weight to make the total loss.

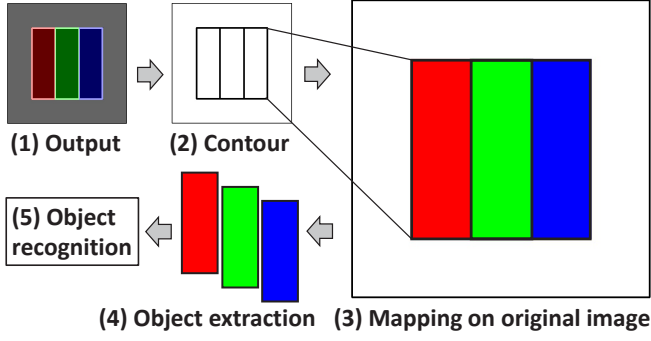


Figure 3: Object detection from contour emphasis image

As mentioned in Sec. 2, in Cycle-Gan training, ground truth data are not used, that is, the error of the extracted contour cannot be monitored. Also, its loss and the error of the object created by it do not always consistent. So, in this method, the difference between the contours, namely their error, shown in Fig. 2 (2-1) and (2-2) is monitored as metrics. The former is the contour emphasized data  $b$ , which is added as ground truth and made from target image (1) by emphasizing the contour of each object; The latter is generated from the corresponding image  $a$  by the generator  $G_{ab}$ , namely the fake image of  $b$ . The training ends when the difference of contours generated from each of them becomes the minimum. Note that such ground truth is not necessary for Cycle-GAN training itself.

I show the error monitoring feature of the CE-model in the dashed rectangle of Fig. 2. The contours of (2-1) and (2-2) are extracted, and blur is applied as shown in (3-1) and (3-2), and the absolute difference between both is made as shown in (4). Lastly, the mean absolute error (MAE) between them is calculated, which is the average brightness of the image shown in (4). This MAE is used for the metrics. Here, blur is for reflecting the distance between both contours into the metrics. For example, when the contours are extremely close, MAE is small and increases as the distance increases.

### 3.2 Object Detection Using Contour Emphasis Image

Figure 3 shows the flow of object detection using the contour emphasis image that is the output of the CE-model. Figure 3 (1) shows the output corresponding to Fig. 2 (2-2). The contour image shown in Fig. 3 (2) is constructed from the image shown in (1) by the procedure of converting this image to grayscale and extracting the area with brightness above the specified threshold as the contour area.

The image in Fig. 3 (2) is reduced because it is the output of the CE-model. Therefore, the contour of each object of this image is converted to its original scale, then mapped to the original image as shown in (3). Based on this contour, as shown in (4), each target object area is extracted. After that, object recognition is performed by the method according to the application such as template matching, multi-class classification of DL, or character recognition.

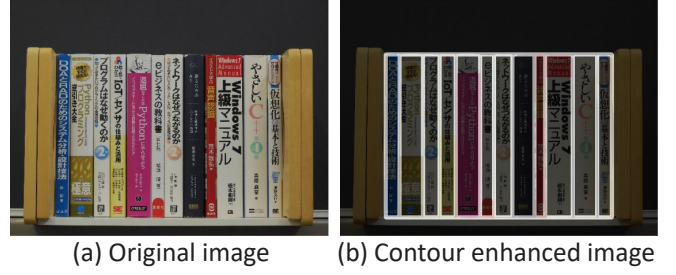


Figure 4: Target image of experiment: books on bookshelf

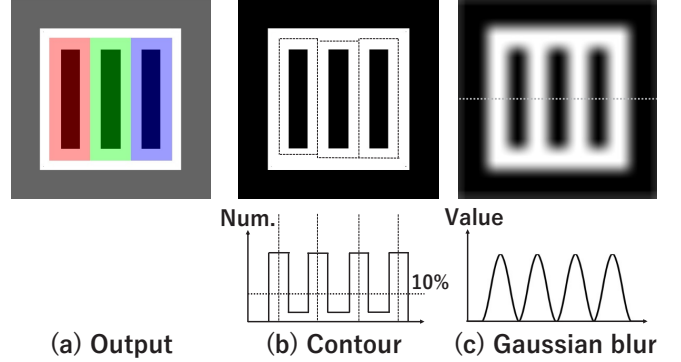


Figure 5: Contour extraction and blur for metrics

## 4 IMPLEMENTATION AND EVALUATIONS

### 4.1 Configuration of Experimental System

To evaluate the effectiveness of the proposed method, I constructed an experimental system for the target's contour extraction and conducted experiments. The books arranged on the bookshelf as shown in Fig. 4 were used for the experimental targets as still images including plural dense objects. Figure 4 (a) shows the original image, and (b) shows the image that emphasizes the contour. Each image size was  $691 \times 518$  pixels; the brightness of each RGB color channel was in the range  $[0, 255]$ ; all the contour was set manually. For (b), to emphasize the contour, the brightness was mapped based on the threshold 128 with an error of 28. That is, the contour area was mapped into the range  $[156, 255]$ , and another area into  $[0, 99]$ .

The experiment was carried out on a personal computer, which CPU was i9-10850K (3.6 GHz), memory was 64 GB, and GPU was GeForce RTX 3090 with 24 GB memory, and OS was Windows 10. Tools and programming languages were Keras Ver. 2.4.3, Tensorflow-GPU Ver. 2.4.1, and Python Ver. 3.7.10. Code of the CE-model was created based on the published Cycle-GAN code [2], along with adding the necessary functions to experiment, such as calculating the metrics, and saving and displaying results.

For the CE-model training, the above bookshelf images were resized to  $128 \times 128$ , and the batch size was 32. For the configuration of this model, residual networks (ResNet) were used, which have shortcut connections for skipping one or more layers to increase the depth of the network [6]. In addition, the weights of the losses in Eqs. (3) to (5) were set



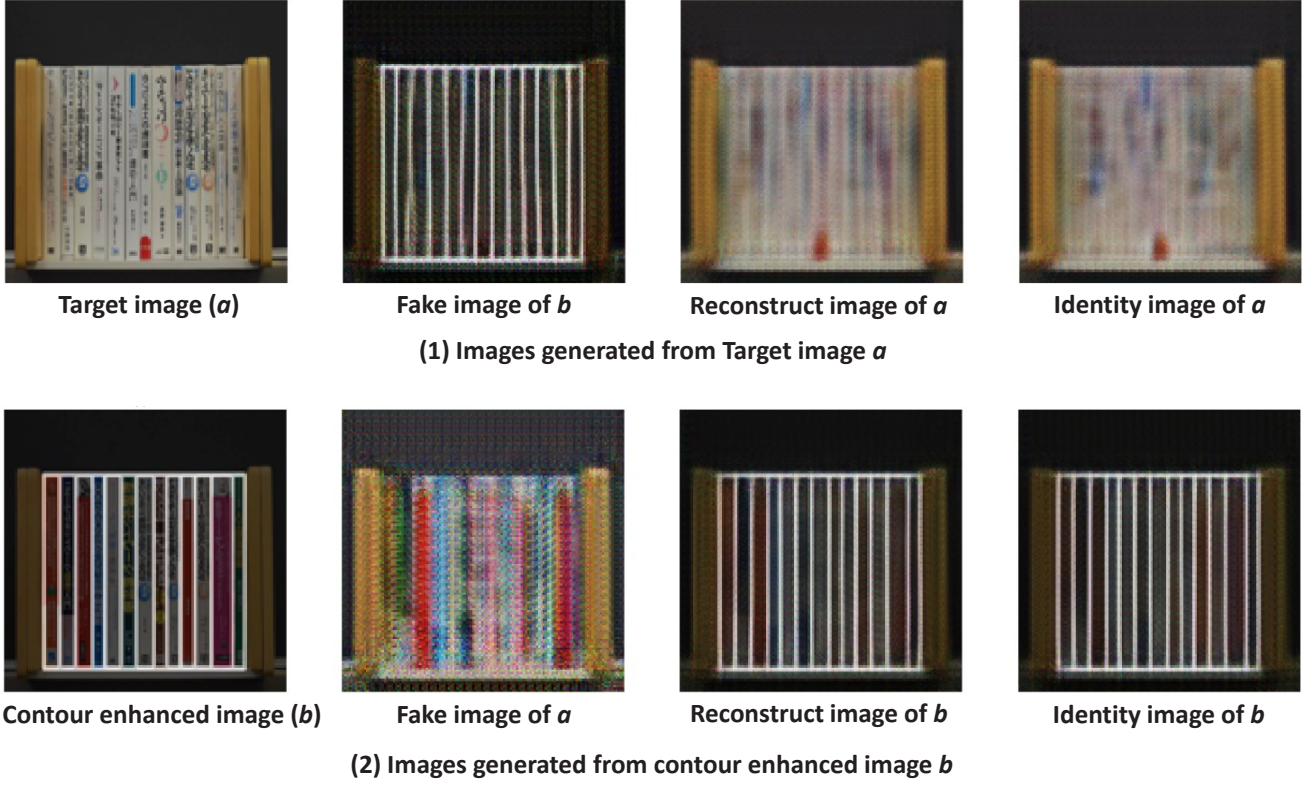


Figure 6: Example of generated images using CE-model

to 4, 10, and 2, and their weighted summation was reflected for the training.

Figure 5 shows the implementation to extract the contour shown in Fig. 3 (2) and to blur it for calculating the metrics as shown in Fig. 2 (3). The image shown in Fig. 5 (a), which was the output image of the CE-model shown in Fig. 2 (2-2), was converted to grayscale. Then, the contour and another area were separated based on a threshold, which was set to 111 in consideration of the grayscale error. As a result, the contour area was set to white, and another area was set to black as shown in Fig. 5 (b). Then, Gaussian blur with kernel size  $5 \times 5$  and standard deviation of 0.3 was applied as mentioned in Sec. 3.1, and the image shown in Fig. 5 (c) was created.

To calculate the metrics, pair of blurred images were used. One was made from the target's contour emphasized image as shown in Fig. 2 (3-1); another was made from the output image as shown (3-2). Then, the MAE of the absolute difference between both, namely metrics, was calculated by Eq. 6.

$$m_e = \sum_{i=1}^n \sum_{j=1}^n |x_{eij} - b_{ij}| / n^2 \quad (6)$$

Here,  $b_{ij}$  and  $x_{eij}$  show brightnesses in pixels of the former and latter respectively;  $e$  is the epoch number;  $i$  and  $j$  indicate the pixel coordinates  $(i, j)$ , and  $n$  is the number of pixels in each coordinate axis. In the training,  $m_e$  was calculated for each epoch. Every time  $m_e$  became the minimum compared before, the model weight was saved as the best weight. Then, when  $m_e$  did not improve the specified number of times, the training was completed. In this experiment, this number was

set to 15 epochs.

Contour extraction was performed on the image in Fig. 5 (b), and firstly the vertical contour was detected. As the procedure, as shown in the lower graph of (b) was created, which indicated the number of pixels in the contour area in the vertical direction for each horizontal position. And, the horizontal area, of which the contour pixel number was 10 % or more was selected as a contour candidate area. In the case of Fig. 5, there were 4 candidate areas, and for each area, the position of the largest number of pixels was selected as the vertical part of the contour, which is indicated by the broken line.

Next, the number of pixels in the contour area of the horizontal direction was counted between each adjacent vertical part of the contour, and the vertical area with 50 % or more numbers was selected as the horizontal part candidates of the contour. And, similar to the vertical part, the horizontal part of the contour was selected. Then, the contour of each object was extracted as shown by the broken line in Fig. 5 (b). In this experiment, they were rectangles.

## 4.2 Evaluations of Contour Extraction

CE-model was trained using the training data, then the accuracy of contour extraction was evaluated. For training and evaluation, 128 pairs of the original image of the bookshelf and the contour emphasis image shown in Fig. 4 were prepared and inflated to 256 pairs by horizontal flip. They were divided into 204 pairs of training data and 52 pairs of test data for monitoring metrics  $m_e$  of Eq. 6. Data were randomly selected and shuffled in each batch of training, without consid-

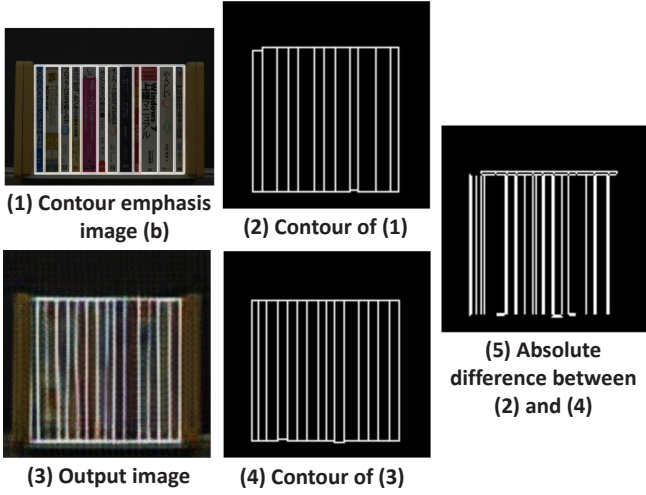


Figure 7: Error of extracted contour

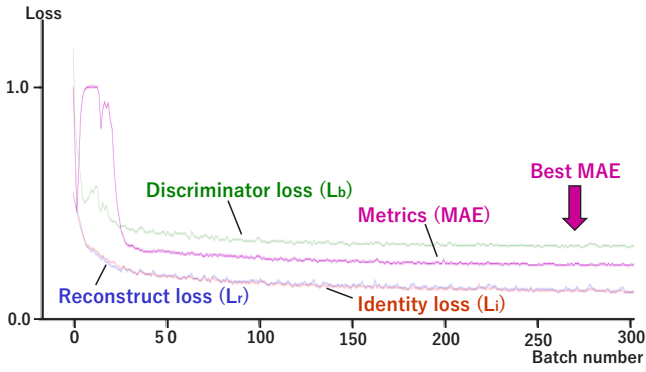


Figure 8: Transition of losses and metrics with training

ering the pair. In this experiment, the best metrics 0.2278 was obtained at epoch 54.

Figure 6 shows examples of images generated by the training data at the above-mentioned epoch 54 during the training. Similar to Fig. 2, the target image, namely the books on the bookshelf, is shown by  $a$ , and the contour emphasis image is shown by  $b$ . The upper row shows the images generated from  $a$ . From the left, the original image  $a$ , the fake image represented by  $G_{ab}(a)$ , the reconstructed image  $G_{ba}(G_{ab}(a))$ , and the Identity image  $G_{ba}(a)$  are shown. Similarly, the lower row shows images generated from  $b$ . Since the pairs were not maintained as above-mentioned, the books are different from the upper row.

As a result, as shown in the second image from the left of Fig 6 (1), clearly emphasized contours were generated from the target image. However, as shown in the right two images in (1), the clear characters in the book's back cover could not be generated in the output of the CE-model.

A contour emphasis image  $b$  of the test data and the contour extracted from it are shown in (1) and (2) of Fig. 7. Similarly, a fake contour emphasis image and the contour are shown in (3) and (4), which were generated from the target image  $a$  corresponding to the above-mentioned  $b$  using the CE-model after training. The aspect ratio of (1) is different because it is before resizing ( $691 \times 518$ ) and the others are output images

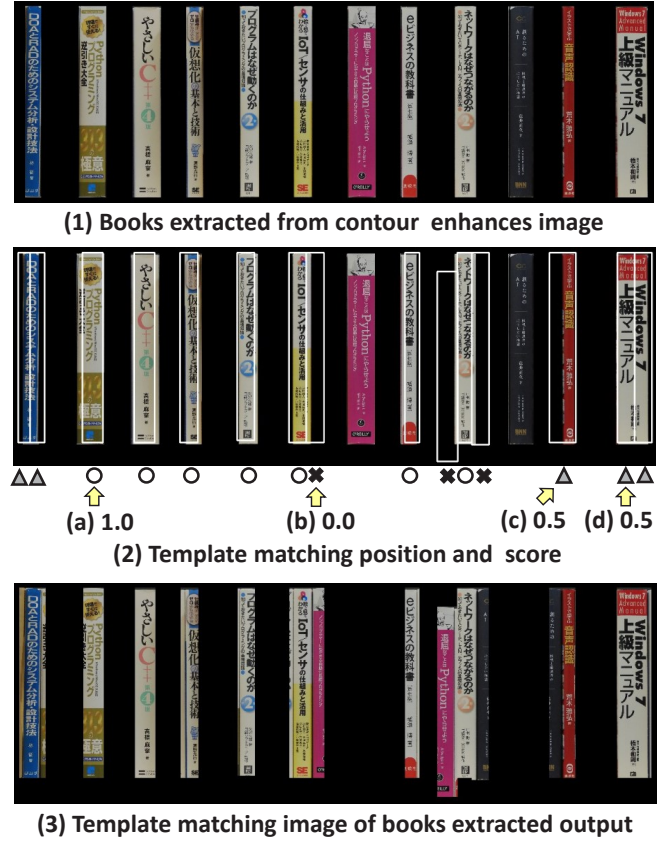


Figure 9: Book recognition results using extracted contours

from the CE-model ( $128 \times 128$ ). Since there was a difference between contours (2) and (4), the absolute difference between them was created as shown in (5). The thin lines indicate that their corresponding contours are in different positions, and the thick lines show the contours adjacent to each other though they are in different positions.

In summary, contours could be extracted using the CE-model, but there were some errors in contour position.

Figure 8 shows the transition of each loss and metrics of the CE-model during training. The horizontal axis of the figure shows the batch number, and 1 epoch corresponds to 5 batches. During the period immediately after the start of training, each loss and metrics showed an individual tendency, but after 30 batches, namely 6 epochs, they showed the same tendency. The arrow indicates epoch 54 (batch number 270) when the metrics (MAE) became the best, namely the minimum.

### 4.3 Evaluations of Object Recognition

To evaluate the effect of contour extraction error described in Sec. 4.2 on object recognition, all test data excluding flipped images were evaluated by template matching. Concretely, for each test data, a pair of contours shown in (2) and (4) of Fig. 7 was used, and each book image was extracted from the original image shown in (a) of Fig. 4 by the procedure shown in Fig. 3.

Figure 9 (1) shows each book image extracted from the original image using the contour shown in Fig. 7 (2), namely contour of the contour emphasis image. Similarly, each book

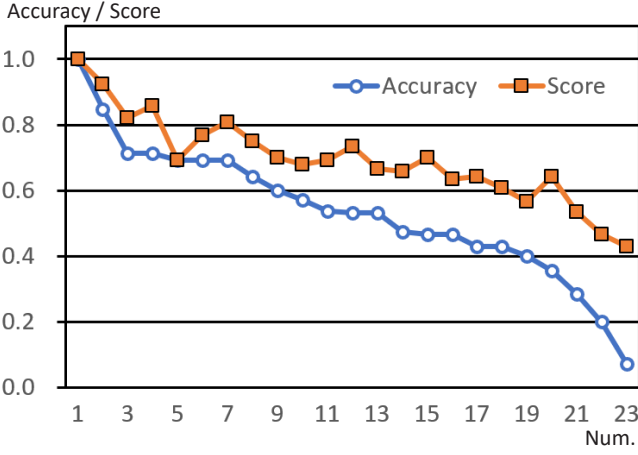


Figure 10: Variation in book recognition accuracy

image was extracted using the contour shown in Fig. 7 (4), namely contours based on the output image of the CE-model. Next, using each latter book image, the corresponding book in the image shown in Fig. 9 (1) was searched using template matching, in which the normalized correlation coefficient matching method was used. (2) shows the matching position of each book by a white rectangle; (3) shows each book image based on the extracted contour shown in Fig. 7 (4).

The following three scores were introduced according to the matching level to evaluate the accuracy of object recognition, as shown in Fig. 9 (2). (a) shows the case of almost the same indicated by  $\bigcirc$ , which score is 1.0; (b) shows the case of almost the difference indicated by  $\times$ , which score is 0.0; (c) and (d) shows the case between the both indicated by  $\triangle$ : there is an extra area or only a part area respectively whereas the same. Note that (c) and (d) are the case when 50% or more of the target book is included, and (b) is the case of less than 50%.

The percentage of the extracted contours with a score of 1.0 (hereinafter, estimation accuracy), and the average score were calculated. In the case of Fig. 9, the number of extracted contours was 15, 1.0 was 7, 0.5 was 5, 0.0 was 3, so the estimation accuracy became  $7/15 = 0.47$  and the average score  $9.5/15 = 0.63$ . Figure 10 shows a graph in which the estimation accuracy and average scores are arranged in descending order of estimation accuracy for all test data excluding flipped images. Both varied widely depending on the type of book, with averages of estimation accuracies 0.537 and of scores 0.694.

The image with the best and worst accuracy are shown in Fig. 11 (1) and (2) respectively. As shown in (1), the best accuracy was obtained when only white books were targeted. On the contrary, as shown in (2), the accuracy of colored books, especially dark-colored books, tended to deteriorate. This tendency was the same in the case shown in Fig. 9, and the darkest-colored book could not be detected. Conversely, white books were relatively well detected.

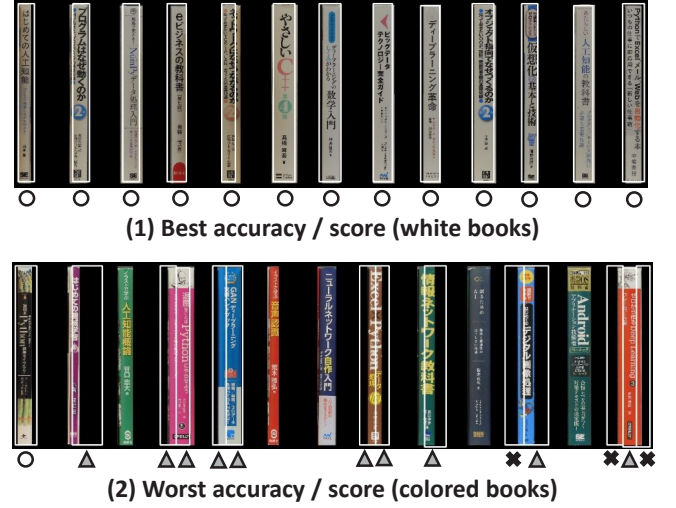


Figure 11: Best and worst results of book recognition using extracted contour

## 5 DISCUSSIONS

The advantage of this method is that it is not necessary to prepare pairs of the original and ground truth images. In this experiment, to monitor the metrics, these pairs were prepared, namely original and contour-emphasized images. However, unlike my previous study mentioned in Sec. 2, for the CE-model in this study, as shown in Fig. 8, the transition of the metrics and CE-model's losses indicated the same tendency. The reason is considered that both are dealing with the image itself. That is, the ratio of metrics data can be reduced. Therefore, for uniform objects such as books on a bookshelf, it is expected to generate training images efficiently. For example, the contour-enhanced image can be divided as shown in Fig 9 (1) and rearranged for augmentations.

In this study, the method to extract contours of the plural dense objects from the still image was proposed, which is based on Cycle-GAN to translate an original image into a contour emphasis image. And, the experiments targeting books on the bookshelf were conducted. As a result, it was shown that their contours could be extracted and each book could be recognized using these contours as shown in Figs. 6 and 9. In addition, it was shown that plural contours could be extracted collectively as shown in Fig. 9.

However, as shown in Fig. 10, the accuracy of extracted contour was dispersed depending on the color of the target book. As shown in Fig. 11, when the book's back cover was white, the contour was extracted almost correctly. On the other hand, the accuracy deteriorated for a dark-colored book. Its reason is considered that the shadow of the boundary of the book greatly contributes to extracting contour, and this improvement is a future challenge.

The advantage of contour detection is considered that it is not necessary to consider the size of the target object for object detection. In general, since the size of the target is unknown in object detection, the approach is adopted, in which several sizes of the template are prepared and applied sequentially. On the other hand, in this method, since the size of the target can be grasped from the extracted contour, for exam-

ple, it is possible to change the size of the template itself in advance of template matching such as shown in Fig. 9.

## 6 CONCLUSIONS

When recognizing small objects in a still image, it is necessary to detect the object first. However, to detect each object in a pixel-by-pixel manner, pairs of the original and ground truth images had to be prepared as training data in the conventional methods using DL, and it caused an obstacle to their practical use.

In this study, I proposed an object detection method based on the contours, in which a contour emphasis image was generated from the original image by the model applying CycleGAN, namely the CE-model. This model is trained mutually using the original and contour emphasis images, and both do not need to associate as such pairs. Furthermore, through the experiments targeting books arranged on a bookshelf, it was shown that it was possible to collectively extract contours and recognize each object using them even from dense objects in a still image.

Future studies will focus on contour extraction accuracy improvement of this model, and its effectiveness evaluations for objects of various shapes.

## ACKNOWLEDGMENTS

This work was supported by JSPS KAKENHI Grant Number 19K11985.

## REFERENCES

- [1] X., Chu, A. Zheng, X. Zhang, and J. Sun, "Detection in crowded scenes: One proposal, multiple predictions," *Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition*, pp. 12214–12223 (2020).
- [2] D. Foster, "Generative deep learning: teaching machines to paint, write, compose, and play," O'Reilly Media (2019), [https://github.com/davidADSP/GDL\\_code](https://github.com/davidADSP/GDL_code) (referred May 24, 2021).
- [3] P. Gao, T. Tian, L. Li, J. Ma., and J. Tian, "DE-CycleGAN: An object enhancement network for weak vehicle detection in satellite images," *IEEE J. Selected Topics in Applied Earth Observations and Remote Sensing*, Vol. 14, pp. 3403–3414 (2021).
- [4] X.Y. Gong, H. Su, D. Xu, et al., "An overview of contour detection approaches," *Int. J. Autom. Comput.*, Vol. 15, No. 6, pp. 656–672 (2018).
- [5] I.J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, and Y. Bengio, "Generative adversarial networks," *Communications of the ACM*, Vol. 63, No. 11, pp.139–144 (2014).
- [6] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 770–778 (2016).
- [7] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask r-cnn," *Proc. IEEE Int. Conf. Computer Vision*, pp. 2961–2969 (2017).
- [8] P. Isola, J.Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 1125–1134 (2017).
- [9] T. Kudo, and R. Takimoto, "CG utilization for creation of regression model training data in deep learning," *Procedia Computer Science*, Vol. 159, pp. 832–841 (2019).
- [10] T. Kudo, "A proposal for article management method using wearable camera," *Procedia Computer Science* Vol. 178, pp. 1338–1347 (2020).
- [11] T. Kudo, "Moving Object Detection Method for Moving Cameras Using Frames Subtraction Corrected by Optical Flow," *Int. J. Informatics Society*, Vol. 13, No. 2, pp. 79–91 (2021).
- [12] T. Kudo, "CG training model application method using cycle-consistent adversarial network," *Int. J. Informatics Society*, Vol. 12, No. 1, pp.41–48 (2020).
- [13] M. Li, Z. Lin, R. Mech, E. Yumer, and D. Ramanan, "Photo-sketching: Inferring contour drawings from images," *2019 IEEE Winter Conf. Applications of Computer Vision (WACV)*, pp. 1403–1412 (2019).
- [14] T.Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 2117–2125 (2017).
- [15] T.Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," *Proc. IEEE Int. Conf. Computer Vision*, pp. 2980–2988 (2017).
- [16] L. Liu, W. Ouyang, X. Wang, P. Fieguth, J. Chen, X. Liu, and M. Pietikäinen, "Deep learning for generic object detection: A survey," *Int. J. Computer Vision*, Vol. 128, No. 2, pp. 261–318 (2020).
- [17] M. Mohammadi, A. Al-Fuqaha, S. Sorour, and M. Guizani, "Deep learning for IoT big data and streaming analytics: A survey," *IEEE Communications Surveys & Tutorials*, Vol. 20, No. 4, pp. 2923–2960 (2018).
- [18] M. Rajput, "YOLOv5 is here! elephant detector training using custom dataset & YOLOV5," <https://towardsdatascience.com/yolo-v5-is-here-b668ce2a4908> (referred May 24, 2021).
- [19] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 779–788 (2016).
- [20] D. Rukhovich, K. Sofiiuk, D. Galeev, O. Barinova, and A. Konushin, "IterDet: Iterative Scheme for Object Detection in Crowded Environments," *arXiv preprint arXiv:2005.05708* (2020).
- [21] K. Saleh, A. Abobakr, M. Attia, J. Iskander, D. Nahavandi, M. Hossny, and S. Nahavandi, "Domain adaptation for vehicle detection from bird's eye view LiDAR point cloud data," *Proc. IEEE/CVF Int. Conf. Computer Vision Workshops*, pp. 3235–3242 (2019).
- [22] V. Sandfort, K. Yan, P. J. Pickhardt, and R.M. Summers, "Data augmentation using generative adversarial networks (CycleGAN) to improve generalizability in CT segmentation tasks," *Scientific reports*, Vol. 9, No. 1, pp. 1–9 (2019).



- [23] W. Shen, X. Wang, Y. Wang, X. Bai, and Z. Zhang, "Deepcontour: A deep convolutional feature learned by positive-sharing loss for contour detection," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 3982–3991 (2015).
- [24] A. Sindel, A. Maier, and V. Christlein, "Art2Contour: Salient Contour Detection in Artworks Using Generative Adversarial Networks," *2020 IEEE Int. Conf. Image Processing (ICIP)*, pp. 788–792 (2020).
- [25] Y. Tian, G. Yang, Z. Wang, E. Li, and Z. Liang, "Detection of apple lesions in orchards based on deep learning methods of cycleGAN and YOLOv3-dense," *J. Sensors*, Vol. 2019, Article 7630926 (2019).
- [26] M. Verhelst, and B. Moons, "Embedded deep neural network processing: Algorithmic and processor techniques bring deep learning to IoT and edge devices," *IEEE Solid-State Circuits Magazine*, Vol. 9, No. 4, pp. 55–65 (2017).
- [27] P. Viola, and M. Jones, "Rapid object detection using a boosted cascade of simple features," *Proc. 2001 IEEE Computer Society Conf. Computer Vision and Pattern Recognition*, Vol. 1, pp. 511–518 (2001).
- [28] H. Yang, Y. Li, X. Yan, and F. Cao, "ContourGAN: Image contour detection with generative adversarial network," *Knowledge-Based Systems*, Vol. 164, pp. 21–28 (2019).
- [29] A. Yilmaz, O. Javed, and M. Shah, "Object tracking: A survey," *ACM Computing Surveys (CSUR)*, Vol. 38, No. 4, Article 13 (2006).
- [30] Z. Zeng, Y. K. Yu, and K. H. Wong, "Adversarial network for edge detection," *Int. Conf. Informatics, Electronics & Vision (ICIEV)*, pp. 19–23 (2018).
- [31] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, "Pyramid scene parsing network," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 2881–2890 (2017).
- [32] J. Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," *Proc. IEEE Int. Conf. Computer Vision*, pp. 2223–2232 (2017).



# Verifications of Influence by Unknown Longer Titles of Work on Robustness of Deep Learning NER

Yukihisa Yonemochi<sup>†</sup> and Michiko Oba<sup>‡</sup>

<sup>†</sup>Graduate School of Future University Hakodate, Japan

<sup>‡</sup>Future University Hakodate, Japan  
{g3119008, michiko}@fun.ac.jp

**Abstract** - The objective of this paper is to highlight the problem of deep learning (DL) named entity recognition (NER) for titles of works. Extracting information from input text is an important task for text interactive interfaces such as chatbots and voice interfaces. In the field of natural language processing, NER is known as an information retrieval for such a requirement. Most of the latest methods for NER utilize deep learning with the highest accuracy. In most cases, only word sequences are used as an input feature as the standard. These methods have a problem in recognizing unknown longer compound words. Longer titles can be found in titles of works such as novel, manga, animation, and movies. We verified this phenomenon using following three aspects. First, we verified how the standard DL NER has the problem of longer titles. Second, assuming that adding lexical features improves performance, we verified its effectiveness. Third, we verified that such longer titles are distributed within the real-world existing titles of works. We here in report the results of the verification and suggest the necessity of considering countermeasures.

**Keywords:** Named Entity Recognition, Deep Learning, Feature Selection

## 1 INTRODUCTION

Extracting proper nouns or unique names such as movie titles from input text is an important part of natural language processing (NLP) for developing interactive systems such as chatbots or voice interfaces. Recognizing the named entity of artifacts, as defined by MUC [1] and IREX[2], is the standard method for extracting proper nouns. This is known as the NER in the NLP area. NER has two tasks: tagging and disambiguation. Tagging is the process of generating tag data for the start and end position of a fragment in a text. Disambiguation involves choosing the correct one from several candidate types.

**Table 1** shows the example text that has the fragment "The Bridge on the River Kwai" that can be both the title of a novel or a movie. B-NOVEL and B-MOVIE indicate the beginning of the title of the novel and the movie, respectively. Likewise, I-NOVEL and I-MOVIE show a range of titles. In addition, O refers to OTHER. This is called the BIO-style label, which is typically used for NER tasks. The NER process must first obtain the correct range, and in this example, the range is 2-7 with the index number. The NER is disambiguated next, choosing between novels and movies. In this example, B-MOVIE and I-MOVIE are the correct choices.

The text at first glance seems to be saying, "I saw the bridge", but it actually says "I saw the movie". Knowledge of the title can help correct this meaning in our brain. Humans can intuitively recognize that the title could be a movie, utilizing the verb "saw".

In NER systems, statistical methods are applied to recognize the range and to choose the correct type. In recent years, some machine learning methods have obtained high NER scores. NLP-progress[3] reported more than 90% accuracy for NER for several datasets.

We encountered the problem that unknown, longer unique names often cause errors in the recognition process. In this context, "unknown" indicates a unique name that was not included in the set of training data, and "longer" means a unique name that has more several words of several types than known name. Such unknown, longer names are commonly seen in novel, manga, cartoon, or movie titles. We refer to such names as UnknownLonger throughout this paper. We can assume that some UnknownLongers will not be extracted correctly in the actual system in case of the following situation. DL is trained by texts including a list of existing titles. It can extract the existing titles from an input text with high accuracy. However, if a new longer title is announced, it is not correctly extracted.

The objective of this study is to argue that when using DL NER in an interactive interface, adding lexical information to the feature is necessary. In this paper, we verify the effectiveness of adding a lexical feature to DL NER for extracting UnknownLonger names from texts. The effectiveness of the technique was measured by experiments that were performed using our original dataset with manga, novel, cartoon, and movie titles. For the purpose of comparison, Japanese and English titles are used because longer titles occur more frequently in Japanese.

## 2 DEEP LEARNING NAMED ENTITY RECOGNITION

Recently, machine learning methods have been utilized for NER. They are machine learning methods suitable for time-series data. The same set of features is used for the NER, regardless of the method. First, the input text needs to be tokenized as a sequence of words or morphemes. Latin-derived languages can be tokenized using the space character, and the Japanese language can be tokenized using the Japanese tokenizer[7]. Simultaneously, label data indicating the tag is prepared, which is referred to as "tagged", "annotated", or "labeled". The sequence of tokens is the explanatory variable

Table 1: Typical NER of a title

index	0	1	2	3	4	5	6	7	8	9
input	I	saw	The	Bridge	on	the	River	Kwai	yesterday	.
Label as Movie	O	O	B-MOVIE	I-MOVIE	I-MOVIE	I-MOVIE	I-MOVIE	I-MOVIE	O	O

Table 2: Adding lexical feature for long title

index	0	1	2	3	4	5	6	7	8	9
input	I	saw	The	Bridge	on	the	River	Kwai	yesterday	.
movie	0	0	1	1	1	1	1	1	0	0
novel	0	0	1	1	1	1	1	1	0	0
Label as Movie	O	O	B-MOVIE	I-MOVIE	I-MOVIE	I-MOVIE	I-MOVIE	I-MOVIE	O	O

and the sequence of labels is the objective variable. In reality, the distributed representation of tokens is used as the input feature for machine learning tools. The method used to generate the distributed representation has a strong influence on the NER result. The data input and labels as movies in **Table 1** shows example label data of machine learning for NER.

Most of the studies have used huge tagged corpora to create a trained model. However, Refs. [8] and [9] both proposed to utilizing Wikidata to create large training datasets. In doing so, they support a wide range of vocabulary but still use the same features and labels.

### 3 DEFINING THE PROBLEM

As mentioned in the second section, standard NER methods use only the surface of word and word sequences as the input feature. This can cause tagging errors in the UnknownLonger titles.

Consider the classification model for the NER, which is trained with training data. The part underscored is the title and must be extracted as an argument for some applications.

- I want to watch Star Wars next week.
- When Harry Potter be released?

The number of words in the titles are both 2, and all of those words are nouns. Using this example, the trained model failed to extract the title name if the title was not in the training data and had a longer title. See the following example:

- I saw The Bridge on the River Kwai yesterday.

The title has six words, and the words are article-noun-preposition-article-noun-noun. This is more complex and longer than the examples mentioned above. This can confuse the classification module. This type of situation poses a problem. Artifact work like manga (comic books), animation (cartoons), novels, and movies are being added daily. However, we cannot train classification models every day, because the computing cost is high.

We verified with datasets which were arbitrarily divided by the number of words in the target title. Shorter titles were used for the training model, and the remaining longer titles were used for testing.

The existence of this problem in NER jobs was verified in the experimental results in Section 6.

## 4 ADDING BOOLEAN FEATURE

To improve the accuracy of DL NER of UnknownLonger titles, we propose injecting vocabulary information into the feature. This method has already been proposed for improving NER for gazetteer[10], adding the feature, which is just a Boolean flag, on the input feature. The flag indicates whether at series of words can be found in the database. In the training data, flags can be generated from the labels. In the test data, the flags can also be generated from labels. In the production input data, the flags can be added by searching the word sequences in the database. **Table 2** shows an example of how flags are added to the feature.

In this example, “The Bridge on the River Kwai” is the title of a movie. The flags can be generated from B-MOVIE and I-MOVIE labels during training and testing. In the production time, the flag must be added by searching the name from a database. More precisely, the flags are added to the tensor matrix after the words are translated into a distributed expression.

## 5 EXPERIMENT

For the first and second verification, we conducted an experiment using our newly prepared test data. The purpose of this experiment is verify that typical NER method has a problem on UnknownLonger and adding lexical feature is effective to improve. The steps of the experiment are:

- (test-1) preparing a typical valid deep learning model,
- (test-2) confirm that the problem actually occurs, and
- (test-3) confirm that it can be improved by adding a lexical feature.

The first verification is that result of (test-1) becomes (test-2) due to the influence of UnknownLonger, and the second verification that (test-2) becomes (test-3) due to the effect of the lexical feature. A more detailed description of the experimental environment is explained in this section.

### 5.1 Test Dataset

As we have a problem with long titles, the dataset was specially designed.

The dataset was generated by combining the following two types of texts: spoken statement patterns and titles.

We chose the titles of mangas, novels, cartoons, and movies as target areas. The spoken statements were created manually with 20 statements for each area. List 1 shows an example of statements in movie contexts.

List 1. Statement patterns

I can't wait until % is released  
Will you go see % next week?  
I need to buy a ticket for %

Titles were collected from each area of Wikidata. Wikidata can be searched using a SPARQL[11] query.

List 2 shows examples of movie titles.

List 2. Examples of movie titles

The Brain That Wouldn't Die  
Puppet Master: The Legacy  
Puppet Master 4

**Table 3** shows the list of items that were used in the SPARQL query to collect titles from Wikidata.

Table 3: List of id and target labels of Wikidata

Type	instance_of(P31)	Japanese	USA
Manga	wd:Q21198342	MANGA	
TV Animation	wd:Q63952888	ANIME	
Written work	wd:Q47461344		NOVEL
Animated series	wd:Q581714		ANIME
Movie	wd:Q111424	MOVIE	MOVIE

Different sets of items were used for both English and Japanese. Movies, animations, and manga are in Japanese. As mangas (comic books) were originally created in Japan, it is obvious TV animation shows and movies will be created from them.

## 5.2 Tools

The experimental environment used the existing components. We chose the BERT [6] Tokenizer because it is state-of-the-art technology for NER. **Table 4** lists the environments and tools used in the experiment.

Table 4: Environment and Tools

Computing environment	Google Colaboratory
Platform	Python3.0
Tool	TensorFlow 2.0
Distributed Representation	BertTokenizer BertJapaneseTokenizer
Classifier	TFBertForTokenClassification
Pretrained model	bert-base-uncased(for English) cl-tohoku/bert-base-japanese-whole-word-masking(for Japanese)
batch size	32

## 5.3 Experimental Steps

The experiments were performed through the following steps.

1. Preparing dataset
2. Prepare standard NER job
3. Test the standard way(test-1)
4. Verify UnknownLonger problem (test-2)
5. Verify the effectiveness of lexical feature (test-3)

## 5.4 Dividing data by the number of words

The set of titles are divided into training data and testing data using the number of words in each title. Designating the number of words in the title by  $N$ , the following were used for the training data for the experiments.

- For English text:  $N < 3$
- For Japanese text:  $N < 4$

We chose these numbers so that the dataset is split 7:3. This indicates that English titles are shorter than Japanese titles.

## 6 RESULTS AND DISCUSSION

**Table 5** presents the experimental results. This section discusses the results. The lines for each language test result for test-1, test-2, and test-3 have already been discussed in Section 5.1.

### 6.1 Preparing the Classification Job

The first lines for each language in Table 5 show the results of test-1, the training job using the standard method randomly dividing the dataset into 70% for training and 30% for testing. The weighted average F1 score was 99.9% for English and 99.4% for Japanese. It shows that the prepared job works for this dataset.

### 6.2 Verifying the Existence of the Problem

The second line for each language shows the results of test-2, the training job using the traditional method of dividing the dataset by the criteria of the number of words in the title. The test data lines with a title with less than three words for English and four words for Japanese were used as training data. The remaining data were used as the testing data. From these results, we can verify that the accuracy of classification decreases when testing only UnknownLonger titles. This difference indicates that Japanese texts are more strongly affected by UnknownLonger titles than English text.

### 6.3 Verifying the Effectiveness of the Lexical Feature

The third line for each language in Table 5 shows the results of test-3, the training job with added lexical feature that divides the dataset according to the criteria of the number of words in the title. The data division criteria were the same as those for test-2. The weighted average F1 score was 99.8% for English and 98.0% for Japanese. This indicates that adding lexical features is effective for this job and the dataset.

Table 5: Experiment results

language	division	count of data		lexical feature	Epoc	F1 score			weighted avg.
		training	testing			NOVEL	ANIME	MOVIE	
English	random 7:3	21264	9114		20	0.997	1.000	1.000	0.999
	train w/N <3	21198	9180		15	0.7987	0.7593	0.8111	0.7889
	train w/N <3	21198	9180	added	24	<b>1.0000</b>	<b>0.9998</b>	<b>0.9959</b>	<b>0.9988</b>
language	division	training	testing	feature	Epoc	MANGA	ANIME	MOVIE	weighted avg.
Japanese	random 7:3	21264	9114		20	0.997	0.998	0.986	0.994
	train w/N <4	21771	8607		6	0.5837	0.6212	0.3045	0.4605
	train w/N <4	21771	8607	added	47	<b>0.9932</b>	<b>0.9820</b>	<b>0.9734</b>	<b>0.9803</b>

Table 6: T-testing titles between Japan and US(P=0.05)

	animation		manga		movie		novel	
	ja-JP	en-US	ja-JP	en-US	ja-JP	en-US	ja-JP	en-US
Mean	3.05	2.74	<b>3.49</b>	2.04	<b>4.00</b>	2.78	2.83	3.17
Variance	5.89	2.57	10.45	1.08	10.76	1.82	7.59	2.71
Count	1225	139	1295	26	6978	65371	192	7153
t	2.00		6.56		31.03		-1.72	
P( $T \leq t$ )	0.023		6.24E-08		7E-199		0.043	

## 7 VERIFYING THE DISTRIBUTION OF LONGER TITLES

As a third verification, we investigated the distribution of the titles.

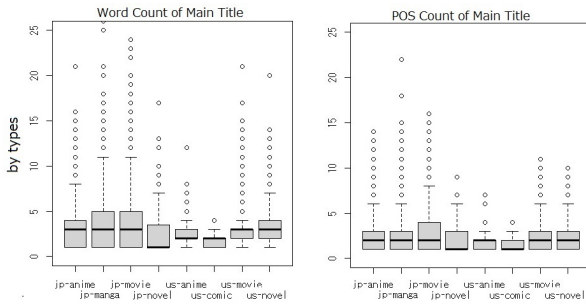


Figure 1: Count of Words and POS of titles by types

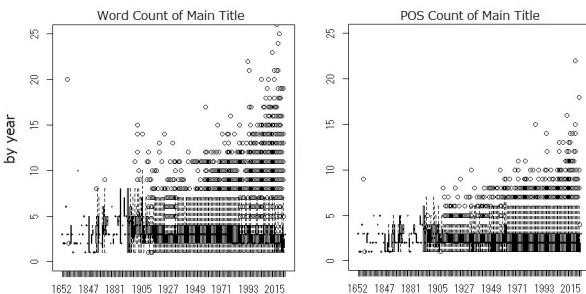


Figure 2: Count of Words and POS of titles by year

**Figure 1** shows the visualization of the distribution of the count of words and parts of speech(POS) by type of work. **Table 6** shows independent t-test result for each type of works. Calculation of the t-test confirmed that Japanese manga and

movie titles are longer than English titles was significant. This confirms that the accuracy of Japanese manga and movie titles in the results in Table 5 was significantly affected. **Figure 2** shows the summaries of the published year. Thus, we verified that the length of titles of work increased by year. This means that UnknownLongers appear after the recognition model is trained.

## 8 CONCLUSION AND FUTURE WORK

We verified that the accuracy of the BERT classifier is affected by the length of the unknown word and can be recovered by adding a lexical feature. From the existing titles of works, Japanese manga and movie titles are longer than English ones. The length of titles of works are increasing every year. Therefore, it is suggested that measures such as adding lexical features are needed to improve the accuracy of identifying UnknownLonger titles. Particular attention should be paid to the Japanese titles of mangas and movies.

On a text interface, people do not talk precise, perfect long names in the input. Abbreviations or shortened names are used. We will study the tendency of people to abbreviate or shorten long title, and devise method to recognize shortened names from input text using lexical information in the future.

## REFERENCES

- [1] N. Chinchor and E. Marsh. Muc-7 information extraction task definition, In Proceeding of the Seventh Message Understanding Conference (MUC-7), Appendices, pp. 359-367 (1998).
- [2] S. Sekine and H. Isahara. Irex: Irie evaluation project in japanese. In Proceedings of the 13th International Conference on Language Resources and Evaluation(LREC), pp. 1977–1980 (2000).
- [3] S. Ruder. Nlp-progress. London (UK): Sebastian Ruder(accessed 2020-01-18). <https://nlpprogress.com> (2020).

- [4] F. Erxleben, M. Günther, Markus Krotzsch, Julian Mendez, and Denny Vrandečić. Introducing wikidata to the linked data web. In *Proceedings of International Semantic Web Conference (ISWC)*, pp. 50–65. (2014).
- [5] ACL SIGNLL. The signll conference on computational natural language learning. <https://conll.org/> (2020).
- [6] J. Devlin, M.-W. Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805* (2018).
- [7] Janome. <https://mocabeta.github.io/janome/> (2020).
- [8] A. L. F. Shanaz and R. G. Ragel. Named entity extraction of wikidata items. In *2019 14th Conference on Industrial and Information Systems (ICIIS)*, pp. 40–45. IEEE, (2019).
- [9] S. Cucerzan. Large-scale named entity disambiguation based on wikipedia data. In *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL)*, pp. 708–716, (2007).
- [10] S. Magnolini, V. Piccioni, V. Balaraman, M. Guerini, and B. Magnini. How to use gazetteers for entity recognition with neural models. In *Proceedings of the 5th Workshop on Semantic Deep Learning (SemDeep-5)*, pp. 40–49 (2019).
- [11] w3c. Sparql query language for rdf. (2008).
- [12] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, et al. Tensorflow: A system for large-scale machine learning. In *12th “USENIX” Symposium on Operating Systems Design and Implementation (“OSDI” 16)*, pp. 265–283 (2016).
- [13] E. Bisong. Google colabatory. In *Building Machine Learning and Deep Learning Models on Google Cloud Platform*, pp. 59–64, (2019).





Session 2:

AI II

( Chair: Hiroshi Mineno )



# Predicting Microclimate Based on Difference from Meteorological Observatory

Genki Nishikawa<sup>†</sup>, Takuya Yoshihiro<sup>‡,\*</sup>

<sup>†</sup>Graduate School of Systems Engineering, Wakayama University, Japan

<sup>‡</sup>Faculty of Systems Engineering, Wakayama University, Japan

\* JST PRESTO, Japan

{s226183, tac}@wakayama-u.ac.jp

**Abstract** - Microclimate means the climate observed near the ground. Observed values of microclimate are useful for agriculture, town management, etc. To measure microclimate such as temperature and humidity, it is necessary to locate sensors at all the observation points on the ground. However, keeping sensors installed at many points is unfeasible due to its large cost. Hence, although some studies use machine learning techniques to learn the effects of topography to predict microclimate, it is difficult to obtain sufficient amount of data to predict microclimate measurement because of a wide variety of factors that effects on microclimate. To solve the problem, there is a study that predicts microclimate measurement at an observed point based on difference between the observation data of microclimate and the observed values of a nearby meteorological observatory, which is computed by a past dataset. However, this method has a problem. This method simply adds the average of the differences between them to the observed values of the meteorological observatory, and thus the prediction accuracy is low. Therefore, in this paper, we propose a method that classifies each prediction time into several weather classes, and make prediction independently for each weather to improve the prediction accuracy.

**Keywords:** IoT, microclimate prediction, town management

## 1 INTRODUCTION

Microclimate is the climate observed in a small area near the ground. Microclimate data is useful in several applications. For example, cultivation management using the observed microclimate in farm field contribute to automation and increases agricultural crops. Also, urban microclimate data is useful to avoid the bad effect on our health such as heatstroke. Consequently, to grasp microclimate is important. Microclimate depends on the surrounding environment and it depends on locations[1]. Hence, although it is necessary to locate sensors at all the observation points on the ground in order to measure microclimate such as temperature and humidity, keeping sensors installed at many points is unfeasible due to its large cost. Therefore, some studies proposed methods to predict microclimate.

Ueyama proposed a method that uses machine learning technique to learn the effects of topography, and predict microclimate based on topographical map[2]. However, it is dif-

ficult to obtain sufficient amount of data to predict microclimate measurement because of a wide variety of factors such as shape, materials and colors of buildings that effect on microclimate as well as anthropogenic heat and plants effect on microclimate.

To solve the problem, Kumagai et al. proposed a method that predicts microclimate measurement at an observed point based on the difference between the observation values at the observed point and those of the nearby meteorological observatory[3]. However, this method has a problem. This method simply adds the average of the differences between them at same time of past days to the observed values of the meteorological observatory, and thus the prediction accuracy is low.

In this study, we focus on the observation that the difference depends on weather. For example, difference between temperature observed on a road paved with asphalt and the temperature observed at meteorological observatory surrounded by lawn will depend on the amount of sunlight. In this paper, we focus on weather as an important factor that increase the prediction accuracy. We propose a method that classifies each prediction time into several weather classes and makes prediction independently for each weather to improve the prediction accuracy.

The remainder of the paper is organized as follows. In section 2, we describe related work. Section 3 describes our method and materials. Section 4 presents the evaluation methods and results with our method. In section 5, we provide conclusions.

## 2 RELATED WORK

### 2.1 Prediction Based on Machine Learning

Suzui et al. proposed a method that predicts microclimate with RNN (Recurrent Neural Network) based on the observed values at the prediction places and nearby meteorological observatory[4]. The method reflects the effect of surrounding environment on the predicted values because it is based on the relation between the measurements at the prediction places and the nearby meteorological observatory. However, we need a large volume of learning data to use machine learning techniques, and observing microclimate measurements at the prediction places for a long time is practically unfeasible.

## 2.2 Prediction Based on Topography Effects

Ueyama et al. proposed a method that predicts microclimate by learning the effects of topography[2]. The method is mainly for farm fields including mountain areas, and predicts the highest and the lowest, and the average temperature of each day based on topography for each square area of the 10 meter grid of a map. In this study, they assume to install environmental sensors at several square cells, and utilize the differences between the observation values of the cells and that from the nearby meteorological observatory. The study computes the difference in temperature according to several features in topography as the explanatory variables to estimate the effects of topography on temperature by means of stepwise regression. Then, they predict temperature for all cells based on the difference. In their evaluation, they installed sensors at 22 places, and predicted the highest and the lowest temperature, and the average of each day in each cell. As a result, the RMSE (Root Mean Square Error) values of the daily average temperature at the observed places was 0.8 °C, that of the highest temperature was 1.4 °C, that of the lowest temperature was 1.2 °C, respectively. Their RMSE might be considered a little large because the RMSE value of the pin-point prediction provided by Japan Meteorological Agency is 1.5 °C.

## 2.3 Predicting Based on Meteorological Observatory

Kumagai et al. proposed a method that first observes microclimate at several prediction places, and predicts the microclimate measurements using the difference between the observed values at the prediction places and that of the nearby meteorological observatory at each predefined time segment [3]. In this study, their method predicts microclimate of each time segment by adding the differences of each time segment to the observed values of the meteorological observatory. Consequently, if we install a sensor at a prediction place for a certain period of time, we can predict the microclimate measurements at the place even after the sensor is removed. However, although the difference depends on weather and surrounding environment, this study do not take it into account. Hence, there is room for improvement.

## 3 PROPOSED METHOD

### 3.1 Overview

In this study, we propose a method that predicts microclimate with higher accuracy than the existing methods by extending the method of Kumagai et al.[3]. We observed the measured data and focused on that the difference between the observation values at the observed points and that of the nearby meteorological observatory has different trends depending on weather. Figure 1 shows the differences between the observation values at the observed points and the meteorological station in Wakayama City. Figure 1 (a) shows an examples of a sunny day. We can see that the difference increases

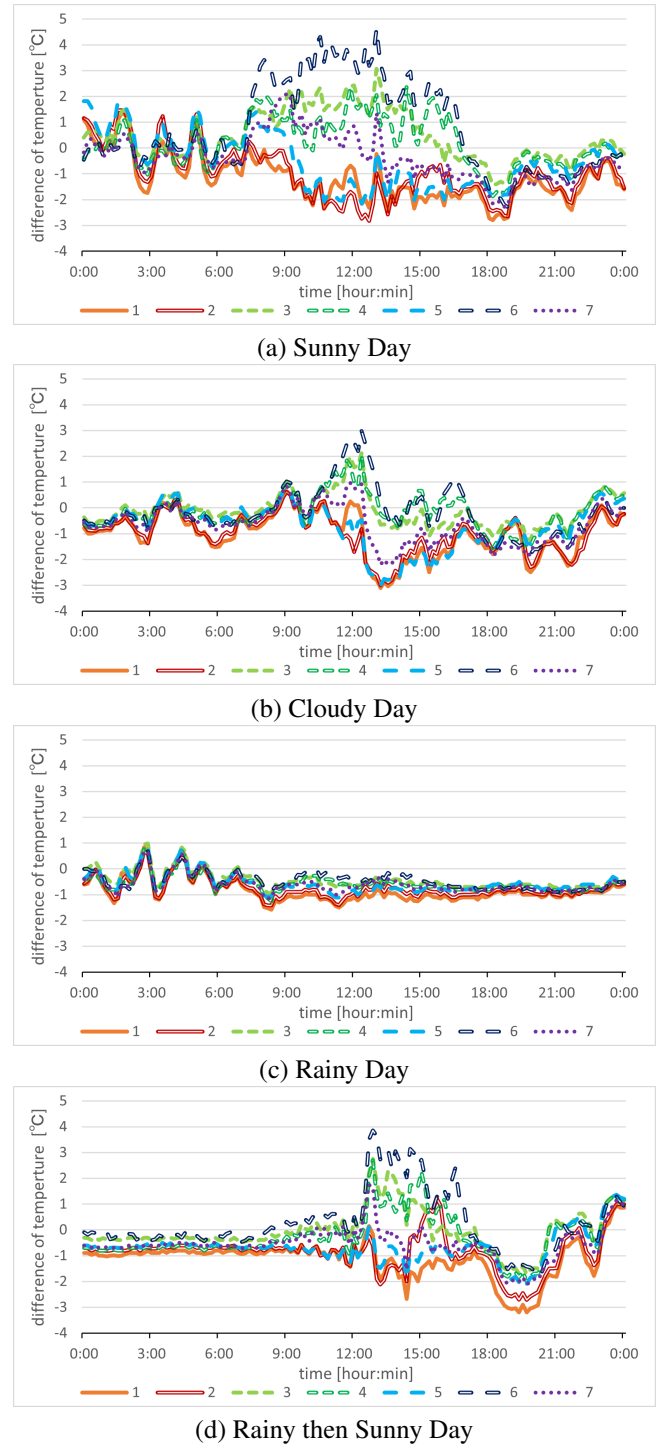


Figure 1: Differences in Measurements According to Weather

as the sun rises. We considered this phenomenon results from that sunlight has small effect on observed values at the meteorological observatory while has large effect on observed values at the points to predict. By contrast, figure 1 (c) shows a rainy day, and the difference is small because of a weak sunlight. In addition, figure 1 (b) shows a cloudy day, and the difference is in between the sunny day and the rainy day. Furthermore, figure 1 (d) shows a rainy weather in the morning and sunny in the afternoon day. According to the meteorological

logical observatory, the time the weather changed to sunny is around noon. Hence, we can see that the difference is small in the morning and it increases rapidly after the time the weather changed.

However, the existing method does not consider that the difference of the observation values has different trends depending on weather.

Consequently, in this study, we propose a method that identifies weather of each prediction time and estimates the differences between the observation values at the observed points and the meteorological observatory at each prediction time independently for each weather to improve the prediction accuracy.

### 3.2 Weather Classification

Our method classifies weather on each prediction time into the three classes, i.e., sunny, cloudy, and rainy. According to Section 3.2, the difference between the observed values at the observed points and that of the nearby meteorological observatory has different trends depending on weather. We classify weather by using the sunlight measurements at the nearby meteorological observatory because we consider that the cause of the difference is mainly the strength of sunlight. Even if the weather is fine, sunlight may be sometimes weak temporarily due to some clouds. Therefore, we use the average value of sunlight strength in the predefined time segment  $P$  before each prediction time  $t$ . We designate the sunlight measurement at time  $t$  by  $l_t$  and classify the weather at time  $t$  by using the average sunlight measurement from time  $t - P$  to time  $t$ . We designate the average value of sunlight measurement by  $L_t = \frac{\sum_{k \in [t-P, P]} l_k}{P}$ , and we classify weather using  $L_t$ . Since there is no sunlight in nighttime, our method classifies weather of nighttime  $t$  using  $L_{t'}$  instead of  $L_t$  where  $t'$  is the sunset time of the previous day. The classification is done based on the following conditions, where  $T_1$  and  $T_2$  are the thresholds to classify, and we determine those threshold values considering the seasonal effect.

sunny day:  $L_t \geq T_1$ .

cloudy day:  $T_1 > L_t \geq T_2$ .

rainy day:  $T_2 > L_t$ .

### 3.3 Microclimate Prediction

Our method installs the environmental sensor and observes microclimate for a certain period of time as the learning data at a prediction place. Afterwards, we remove the sensor and predict the microclimate measurements at the place even after the sensor is removed. Our method classifies weather of each prediction time into the three classes: sunny, cloudy, and rainy. We designate the set of prediction time  $t$  classified into class  $X$  by  $Time_X$  and the measurement at the prediction place  $p$  at time  $t$  by  $v_p^t$ .

$$S_{p,X}^t = \frac{\sum_{t \in Time_X} (v_p^t - V^t)}{|Time_X|}. \quad (1)$$

Here, we designate the average of the difference between the measurement at the prediction place  $p$  and that of nearby meteorological observatory at time  $t$  by  $S_{p,X}^t$ , and obtain  $S_{p,X}^t$  from equation (1) from the differences at time  $t$  of past days we observed microclimate for. Here,  $V^t$  is the measurement value at the meteorological observatory at time  $t$ , and  $|Time_X|$  is the number of time  $t$  of past days, i.e., the number of measurements, that  $Time_X$  includes. In addition, time  $t$  is assumed as a discrete value because the sensor measurements are periodically done in this study.

Our method predicts the unknown measurements at the prediction place  $p$  by using this value  $S_{p,X}^t$ . Let the time to predict be  $t'$ . If we obtain the measurement of the meteorological observatory  $V^{t'}$ , we obtain the prediction value  $e_p^{t'}$  on  $t'$  at  $p$  from the following equation (2).

$$e_p^{t'} = V^{t'} + S_{p,X}^{t'}. \quad (2)$$

## 4 RESULTS

### 4.1 Evaluation Method

We installed sensors[5] and observed microclimate at the 7 prediction places around a building in Wakayama University for 62 days from February 4 to April 5, 2021. Afterwards, we predicted the temperature as a microclimate measurement to compare RMSE between our method and the method without classification[3]. Figure 2 shows the 7 prediction places the map around the building in Wakayama University. The nearby meteorological observatory is Wakayama Local Meteorological Observatory. We put the sensors in the solar radiation shields to shelter the sensors from the direct effect of sunlight, wind, and rain as shown in figure 3. According to the Japan Meteorological Agency's guidebook[6], in meteorological observation, we must install the sensors 1.5 m above the ground to avoid the effect from the ground to the sensors. The guidebook is not for microclimate; however, we installed each sensor 1.5 m above the ground with a tripod because the effect of the ground is too strong in temperature measurement.

The measurement time interval of the sensors were 1 minute. Because the measured values included fluctuations in short period as measurement error, we took the moving average with 10-minutes window to remove the fluctuations. Note that the measurement time interval of the meteorological observatory is 10 minutes also.

In this paper, we used the sunlight measured by the sensors at the prediction places around Wakayama University because Wakayama Local Meteorological Observatory do not measure the sunlight. Although the sensors were in the solar radiation shields, the solar radiation shields can not completely shelter the sensors from sunlight, and we got relative

measurement of sunlight strength. Hence, we can classify the weather according to predict. However, when the sensors were in the shade of the buildings, correct sunlight strength is not obtained. Therefore, we used the largest sunlight measurement among the 7 prediction places.

We used leave-one-out cross-validation in which we predict a microclimate measurement from the dataset excluding the measurement of the predicting day. as a result, we predict microclimate measurements using the measurements of 61 of the prediction place and the meteorological observatory at other 61 days in the 62 days of measurement period.

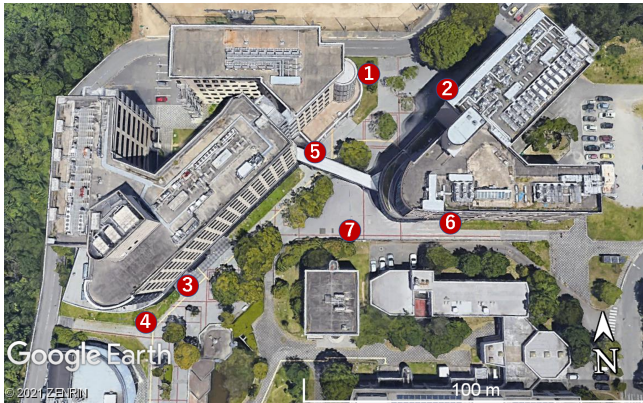


Figure 2: Observed Points ( Source: Google Earth )

## 4.2 Evaluation Results

Table 1 shows the RMSE of the predicted values with our method and the conventional method. In addition to the RMSE values, Table 1 shows the average of those RMSE values for each method. The average of our method is 0.076 °C smaller than average of another method. The difference is about 8.24 percent of the larger RMSE value. Figure 4 shows the RMSE of those values in each time segment from 0:10 to 24:00. The RMSE of our method is smaller than the conventional method in daytime, whereas that of nighttime is almost the same.

Next, we focus on location of sensors show in Figure 2, and compare the RMSE values at each prediction place between methods. See Table 1 again. The RMSE values of both methods at places 1, 2, 5 by both methods are almost the same, and the RMSE values of our method at places 3, 4, 6, 7 are smaller than those of the conventional method. The prediction places 1, 2, 5 are in the shade of buildings or trees in considerable time of a day. Therefore, we found that our method improves the prediction accuracy in the case where sunlight has large effect on the place. From this, it is considered that taking the shade effect into will account improve prediction accuracy.



Figure 3: Observation Device

Table 1: RMSE

Place	RMSE(Proposed)	RMSE(Conventional)
1	0.897	0.897
2	0.791	0.808
3	0.841	0.966
4	0.830	0.931
5	0.904	0.897
6	0.858	1.105
7	0.800	0.848
Average	0.846	0.922

## 5 Discussion

From the results, we discuss about some possibility to improve further the prediction accuracy of the temperature. First, from the results, we see that the amount of sunlight rather than weather make effects on the prediction accuracy. For example, we confirmed from the results that the shadow made by buildings had an effect. To consider the shadow effect, we have to recognize the time of shadows at each places, and reflect it on the prediction method. Furthermore, by observing the results carefully, we found that the strength of sunlight varies even in a day, which is likely to effect on the prediction accuracy. Thus, to improve the prediction accuracy, estimating the effect level as a continuous value rather than categorized weather from the time series of the sunlight strength, and predict the temperature using the level, would be one ef-



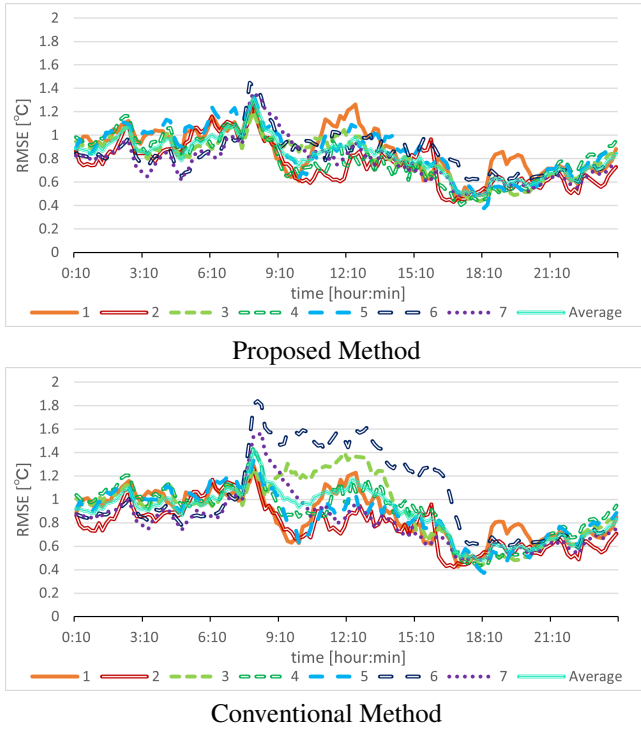


Figure 4: RMSE in time Series

fective strategy for the future.

From careful observation of the prediction results, we also found the prediction error in nighttime is likely to depend on clouds in the sky, i.e., the effect of radiational cooling. If the amount of cloud in the sky can be obtained, this can be a useful source of data for temperature prediction.

## 6 CONCLUSIONS

In this study, we propose a method that classifies weather of each prediction time by using the sunlight and estimates the differences between the observation values at the observed points and the meteorological observatory independently for each weather to improve the prediction accuracy. We measured microclimate at 7 prediction places around a building in Wakayama University for 62 days from February 4 to April 5, 2021 and observed the difference between the measurement and that of the nearby meteorological observatory.

We predicted the temperature as a microclimate measurement to compare RMSE between our method and the conventional method without classification. We calculated the RMSE of the predicted values with our method and the conventional method for each predicted place and the average of those RMSE values for each prediction method. The average of our method is 0.076 °C smaller than average of another method. The difference is about 8.24 percent of the larger RMSE value. In addition, we calculated the RMSE of those values in each time segment from 0:10 to 24:00, and the RMSE of our method is smaller than the conventional method in daytime, whereas that of nighttime is almost the same. Next, we focus on location of sensors, and compare the RMSE values at each prediction place between methods. As

the result, we found that our method improves the prediction accuracy in the case where sunlight has large effect on the place.

From the results, we discuss about some possibility to improve further the prediction accuracy of the temperature. First, from the results, we see that the amount of sunlight rather than weather make effects on the prediction accuracy. In addition, by observing the results carefully, we found that the strength of sunlight varies even in a day, which is likely to effect on the prediction accuracy. Thus, to improve the prediction accuracy, estimating the effect level as a continuous value rather than categorized weather from the time series of the sunlight strength, and predict the temperature using the level, would be one effective strategy for the future. From careful observation of the prediction results, we also found the prediction error in nighttime is likely to depend on clouds in the sky, i.e., the effect of radiational cooling. If the amount of cloud in the sky can be obtained, this can be a useful source of data for temperature prediction.

## ACKNOWLEDGMENTS

This work was supported by JST, PRESTO Grant Number JPMJPR1939, Japan.

## REFERENCES

- [1] Ministry of the Environment, Heat Illness Prevention Information, <https://www.wbgt.env.go.jp/wbgt.php> Referred in May, 2021 (in Japanese).
- [2] Hideki Ueyama, "Developing Applications to Create 50m-mesh Data of Temperature, Sunlight Strength, Relative Humidity, Reference Evaporation", Bulletin of the NARO Agricultural Research for Western Region, No.19, pp. 13-43, (2019) (in Japanese).
- [3] Kenta Kumagai, Toshihira Uchibayashi, Toru Abe, Takuo Suganuma, "Predicting Microclimate from Sensor Data for Town Management", IPSJ SIG Technical Report, , Vol. 2016-IS-138 No.10, pp. 1-7, (2016) (in Japanese).
- [4] Seiji Suzuki, Nahomi Fujiki, "Prediction and Supplement of Meteorological Data Using Recurrent Neural Networks", The 82nd National Convention of IPSJ, pp. 285-286, (2020), (in Japanese).
- [5] Omron, 2JCIE-BL01 Environment Sensor, <https://www.omron.co.jp/ecb/product-info/sensor/iot-sensor/environmental-sensor> Referred in May, 2021 (in Japanese).
- [6] Japan Meteorological Agency, "A Guidebook for Meteorological Observation", [https://www.jma.go.jp/jma/kishou/known/kansoku\\_guide/guidebook.pdf](https://www.jma.go.jp/jma/kishou/known/kansoku_guide/guidebook.pdf) Referred in June, 2021 (in Japanese).





# Estimating the best time to see cherry blossoms using SNS and time-series forecasting of tweet numbers using machine learning

Tomonari Horikawa<sup>\*</sup>, Munenori Takahashi<sup>\*</sup>, Masaki Endo<sup>\*\*</sup>,  
Shigeyoshi Ohno<sup>\*\*</sup>, Masaharu Hirota<sup>\*\*\*</sup>, and Hiroshi Ishikawa<sup>\*\*\*\*</sup>

<sup>\*</sup> Electronic Information Course, Polytechnic University, Japan

<sup>\*\*</sup> Division of Core, Polytechnic University, Japan

<sup>\*\*\*</sup> Faculty of Informatics, Okayama University, Japan

<sup>\*\*\*\*</sup> Graduate School of System Design, Tokyo Metropolitan University, Japan

{b18321,m20305,endou,ohno}@uitec.ac.jp

hirota@mis.ous.ac.jp

ishikawa-hiroshi@tmu.ac.jp

*Abstract* – Collection of tourism information using the web has become popular in recent years. Moreover, tourists are increasingly using the web to obtain tourist information. Particularly because of the spread of social network services (SNSs), various tourism information has become available. Numerous studies have been conducted using Twitter, an SNS. A low-cost moving average method using location information associated with geotagged tweets has been proposed to estimate the optimal time (peak period) for biophenological observations. Geotagged tweets are also useful as a social sensor for estimation and acquisition of local tourist information in real time because the information can reflect real-world situations. We have been working at estimating the best times to view cherry blossoms. Results of earlier studies show that, using weighted moving averages, one can estimate the best time to see cherry blossoms in each prefecture. Nevertheless, it was difficult to estimate the best time to see each tourist spot. Therefore, we considered a method that can estimate the best time to see the cherry blossoms at each tourist spot, considering the weather information. We also demonstrated machine learning for use in predicting the number of tweets in a certain period.

*Keywords:* Machine learning, Mining, Sightseeing, SNS

## 1. INTRODUCTION

In recent years, opportunities for tourists to obtain tourism information using the web are increasing. Particularly because of the spread of social networking services (SNSs), information of diverse kinds is distributed and accumulated on the web. Some SNSs, such as Twitter, can use location information [1]. Our research is examining estimation of the best time to view cherry blossoms using information related to Twitter's geotagged tweets. A low-cost method was proposed [2]: it uses a moving average and this information for estimating optimal times for observing phenology. The proposed method can estimate the best time to view cherry blossoms in prefectures and municipalities where a certain number of tweets with geotags are visible. In addition, the geotagged tweets used for this method are useful as a social sensor to assess real-world situations. Therefore, this effective method of estimating viewing times

can provide local tourism information in real time. As demonstrated by a study using a weighted moving average to estimate the best time to see cherry blossoms [3], it became possible to estimate the best time to see cherry blossoms on a prefecture-by-prefecture basis. However, when estimating the best time to see cherry blossoms at each tourist spot, a problem arose that the estimation accuracy was reduced considerably because of the influence of weather. Therefore, we examined whether the accuracy of the best time to estimate the cherry blossoms at each tourist spot could be improved by adding weather information: a difficulty posed by the existing method. Furthermore, results showed that differences in the amounts of tweets of each year were large. Therefore, in addition to the weather information, we examined the optimization of conditions for estimating the best time to view blossoms. We also investigated the possibility of using machine learning for time-series forecasting of tweet numbers as a new method to estimate the best time to see cherry blossoms. This method uses tweet information accumulated to date as learning data for time-series analysis. Using machine learning, the method shows the possibility of automatically predicting the number of tweets in a certain period. This study investigated the best method of time estimation improvement for the existing method and the time series forecasting method of the number of tweets using machine learning.

## 2. RELATED RESEARCH

Diverse information such as location information, images, and character strings are accumulated by SNSs continuously and in large amounts. Earlier research efforts have examined extraction of such information from SNSs. Hereinafter, we describe research that has examined information extraction from SNSs.

Maenaka et al. [4] proposed the Sakura Sensor, a participatory sensing system that extracts landscape route information automatically from videos taken using an in-vehicle smartphone. It then shares data among users nearly in real time. Using the method described by Maenaka et al.,

we were able to confirm cherry blossoms in a flowering state with accuracy of about 74% and a recall rate of 84%. In addition, the k-stage sensing method achieved the same point of interest detection rate in half the sensing time as that shown by the conventional method.

Amati et al. [5] conducted a temporal analysis of the Twitter stream to investigate the evolution of unique events based on the burst of popularity of associated hashtags. They derived classification of events according to different patterns corresponding to the peak of the volume of exchanged message and to propagation of these events on social networks having characteristics identical to those of Twitter.

Yang et al. [6] describe TimeSeries AggregatoR (TSAR), a robust, scalable, real-time event time series aggregation framework built primarily for engagement monitoring: aggregation of interactions with Tweets, segmentation along a multitude of dimensions such as devices and engagement types. TSAR was built on top of Summingbird to manage the ingestion and processing of events to the publication of results in heterogeneous data stores. Clients were provided a query interface that powers dashboards and which supports downstream ad hoc analytics.

Finally, although research on SNS and time series analysis has been conducted as described above, no study of estimated best times to view cherry blossoms has been reported. Therefore, for this study, we detect the estimated best times to view cherry blossoms.

### 3. EXPERIMENT METHOD

This section presents descriptions of preprocessing and the data used. Using the Streaming API [7], we collected geotagged tweets with location information including latitude and longitude in Japan. Then we analyzed tweets with biological names. For this experiment, conducted during Feb. 17, 2015 through Jul. 29, 2019, the transition of tweets related to cherry blossoms was confirmed with analyzed organism names as "さくら", "サクラ", and "桜". In addition, from latitude and longitude information of the geotagged tweet including location information, the simple reverse geocoding service [8] of the National Research Institute for Agricultural Sciences and Food Industry was used for latitude and longitude information, and for the prefecture or municipality from which each tweet was sent. Analyses were conducted using general towns and streets within the named city planning area.

#### 3.1. Judgment of cherry blossom viewing time

As the best time estimation method, we use the existing method, which uses a weighted moving average. This section presents a description of existing methods.

#### 3.2. Weighted moving average

The weighted moving average used for the existing method is a moving average with each value assigned a weight. By adding weights, the recall and accuracy of the best time estimation are improved. Using the existing method, the median was set to 1. The values  $\pm 0.5$  from the median were taken respectively as the minimum and maximum values. In addition, except for the median, weights from the lowest value to the highest value were assigned linearly. The third decimal place was rounded. Taking the 5-day weighted moving average used in the existing method as an example, the following equation (1) was obtained. Here,  $H_{avg5}$  represents the weighted moving average for 5 days;  $x_y$  denotes the number of tweets  $x$  before  $y$  days prior.

$$H_{avg5} = (x_5 \times 0.5 + x_4 \times 0.75 + x_3 \times 1 + x_2 \times 1.25 + x_1 \times 1.5) / 5 \quad (1)$$

#### 3.3. Best time estimation method

Using a simple moving average and a weighted moving average, the following estimation method was set for the frequency of appearance of each geotagged tweet containing the target word. Results were analyzed by date to estimate the best viewing period. 1. We used a one-year simple moving average to ascertain the period during which tweets about cherry blossoms increased. 2. Because the number of tweets tends to be higher on Saturdays and Sundays, the 7-day weighted moving average was used on a weekly basis. 3. A 5-day weighted moving average was used based on the average number of days from flowering of cherry blossoms to full bloom: 5 days. The best time to view blossoms at each tourist spot was inferred using these best time estimation criteria.

### 4. PROPOSED METHOD

For this study, we propose a method that can estimate the best time to see cherry blossoms at each tourist spot. Earlier studies demonstrated that if a person can see the number of geotagged tweets in a prefecture or city, then that person can estimate the best time to visit. It is noteworthy that the full-bloom day is a state in which about 80% or more buds of the sample tree are open. Therefore, the best time estimate continues even after the full-bloom day. Using this conventional method, one can estimate the best time for viewing after the day of full bloom for a prefecture or municipality unit. However, the existing method presents a difficulty by which the accuracy of the 2018 model is markedly lower because of the influence of the weather. Therefore, we examined whether the accuracy of the best time to estimate the cherry blossoms at each tourist spot could be improved by adding weather information, which was a difficulty posed by the existing method. A large difference was also found to exist in the number of tweets of

each year. Therefore, in addition to the weather information, we also examined the optimization of conditions for estimating the best time to view phenomena. The one-year simple moving average has been changed as an optimization of conditions for estimating the best time to see. Furthermore, we investigated the possibility of using machine learning to predict time-series forecasting of tweet numbers as a new method for estimating the best time to view cherry blossoms.

#### 4.1. Interpolation method

In this section, we propose a method for estimating the best time to see cherry blossoms, which incorporates weather information. During 2017–2019, we surveyed the weather information of Tokyo using goo weather [9]. As a judgment of the weather of the day, the weather at 12:00 on that day was taken as the weather of the day. When we compared weather information with the number of tweets about cherry blossoms in Tokyo, we identified a decrease in the number of tweets about bad weather (rain and snow) days. Therefore, we propose a method to interpolate the number of tweets on bad weather days. As an interpolation method, the regression line was calculated from the number of tweets 2–6 days before the bad weather day. Using the obtained regression line, we interpolated the number of tweets on the prior day, which had bad weather. The number of interpolated tweets is rounded down to the nearest whole number. If bad weather continued, then the interpolation result was updated as necessary to estimate the best time to view the cherry blossoms. When the interpolation result was less than the original number of tweets, the original number of tweets was used without interpolation.

#### 4.2. Optimizing conditions for estimating the best time for viewing

From Feb. 1 through Apr. 30, 2017–2019, the total numbers of tweets about cherry blossoms in Meguro Ward were 2534 in 2017, 1071 in 2018, and 957 in 2019. The one-year moving average was changed in 2018 and 2019 using the ratio. The left side is a comparison of the one-year moving average. The right side is a comparison of the total number of tweets from Feb. 1 to  $y$  day. The formula was transformed. Calculations were performed. The formula for calculation is shown below. Therein,  $X$  represents the one-year moving average of the  $y$  day of the desired year. In addition,  $a$  is the one-year moving average of the  $y$  day of the prior year;  $b$  is the total number of tweets from Feb. 1 to  $y$  day. In addition,  $c$  is the total number of tweets from Feb. 1 of the prior year to  $y$  day.

$$X = \frac{a \times b}{c} \quad (2)$$

#### 4.3. Time-series forecasting method for the number of tweets using machine learning

We will investigate the possibility of using machine learning as a new method for estimating the best time to see cherry blossoms. In the existing method, the best time to estimate is estimated by application of conditions of the best time (1-year moving average, 5-day weighted moving average, 7-day weighted moving average, etc.) to the number of tweets. Therefore, by predicting the number of tweets, it might be possible to estimate the best time to see a set period. For the proposed method, Amazon Web Services (AWS) [10] was used. Time series forecasting was performed using the Amazon Forecast function of AWS. The location, date, and number of tweets were used as learning data. The learning data include tweet information from Feb. 1 through Apr. 30 from 2015–2018 (3 months  $\times$  4 years) and from Feb. 1 through Feb. 28 in 2019 (Total of learning data: about 1 year and 1 month). As described above, the tweet information used as learning data is actually a discontinuous value. However, to perform time series forecasting using Amazon Forecast, it is necessary to make the learning data a continuous value. Table 1 shows that the learning data were stored in a three-column CSV file.

Table 1: Some learning data stored in a three-column CSV file

A	2019-01-27	118
A	2019-01-28	117
A	2019-01-29	94
A	2019-01-30	86
A	2019-01-31	83

In the left column, alphabets that determine the location (A, B, C, ... etc.) are shown. In the middle column, the dummy date (format: yyyy-mm-dd) calculated back from the date you want to predict is stored. In the right column, the number of tweets of learning data is shown. The Amazon Forecast algorithm used DeepAR +, which predicts the number of tweets. The period to be predicted is the two months from Mar. 1 through Apr. 30, 2019. For the predicted value, a value that satisfies the demand of 50% obtained using the weighted quantile loss was used. Weighted quantile loss is a type of metric for forecasting using Amazon Forecast. The predicted number of tweets has been rounded down to the nearest whole number. In addition, when the predicted number of tweets became negative, it was treated as 0 tweets. With the existing method, the best time to see is estimated using the actual number of tweets up to the day before estimating the best time. Therefore, the best time for viewing is estimated every day. However, the proposed method predicts the number of tweets in two months, in advance. We estimate the best time to see using the predicted number of tweets. Therefore, it is possible to estimate the best time to see the two months from Mar. 1 to Apr. 30. Using machine learning, we verified whether it is possible to predict the number of tweets in a certain period automatically.

## 5. EXPERIMENT RESULTS

This chapter presents results obtained for the best time to estimate the cherry blossoms. For this study, the recall and precision values are evaluated to assess the accuracy of estimating the best time for viewing cherry blossoms. The correct answer data use the cherry blossom date to full bloom date observed by the Japan Meteorological Agency. However, the existing method also estimates the best time for viewing after the full bloom date observed by the Japan Meteorological Agency. Therefore, precision can be considered to decrease to some degree.

### 5.1. Best time estimation

Figure 1 shows the best time to see the cherry blossoms in Meguro Ward in 2018, as estimated using the existing method. Figure 2 shows the best time to see the cherry blossoms in Meguro Ward in 2018, as estimated using the interpolation method. Table 2 shows the recall and precision of Meguro Ward estimated by the existing method. Table 3 shows the recall and precision of Meguro Ward estimated using the interpolation method.

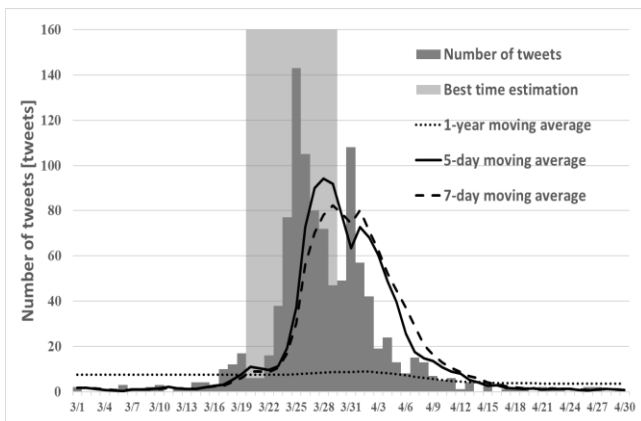


Figure 1: Meguro Ward in 2018.

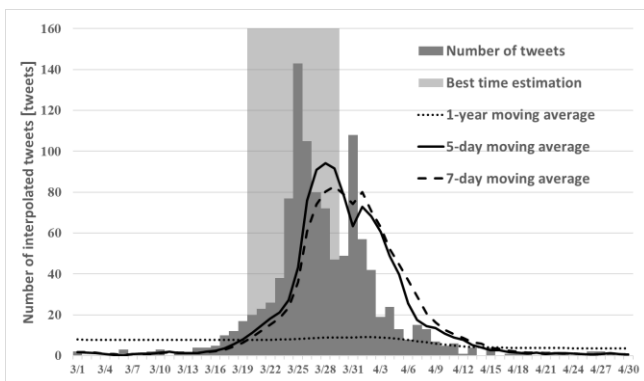


Figure 2: Meguro Ward in 2018 using interpolation.

Table 2: Recall rate and precision rate for cherry blossom viewing (Meguro Ward)

	Recall	Precision
2017	100%	61.9%
2018	62.5%	38.5%
2019	100.0%	41.2%

Table 3: Recall rate and precision rate for cherry blossom viewing using interpolation (Meguro Ward)

	Recall	Precision
2017	100%	61.9%
2018	62.5%	38.5%
2019	100.0%	41.2%

Comparison of Fig. 1 and Fig. 2 shows that the number of tweets changed particularly from Mar. 19 through Mar. 24. We were able to draw a graph that was less affected by bad weather. However, a comparison of Table 1 and Table 2 shows that no change was found in either recall or precision.

Therefore, the estimation was performed next using interpolation and optimization of conditions for estimating the best time for viewing. Figures 3 and 4 present results obtained for cherry blossoms in Meguro Ward in 2018 and 2019.

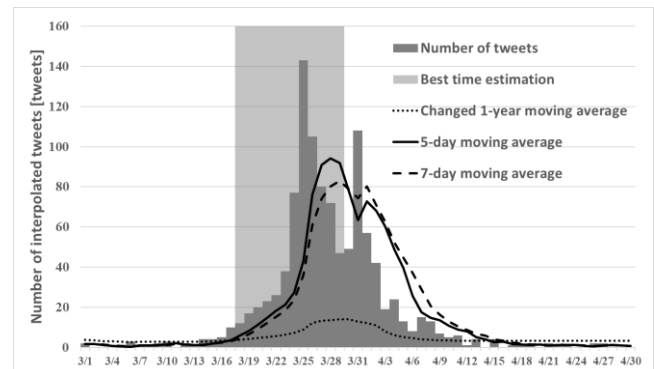


Figure 3: Meguro Ward in 2018 using interpolation and optimization of conditions.

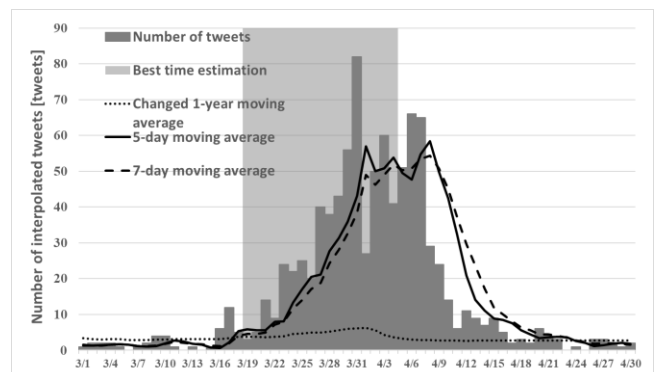


Figure 4: Meguro Ward in 2019 using interpolation and optimization of conditions.

Next, Table 4 compares the time from flowering to full bloom of the cherry blossoms in Tokyo observed by the Japan Meteorological Agency and the estimated the best time to see cherry blossoms in Meguro Ward by the interpolation method and optimizing of conditions. The period from flowering to full bloom of the Japan Meteorological Agency is indicated by a black arrow. The gray part represents the best estimated time to view the blossoms. Table 5 shows the recall and precision rates of the estimated cherry blossoms in Meguro Ward in 2018 and 2019 estimated using optimization of conditions and interpolation for the number of tweets on bad weather days.

Table 4: Comparison of cherry blossom viewing times

2018	Estimated	2019	Estimated
3/10		3/10	
3/11		3/11	
3/12		3/12	
3/13		3/13	
3/14		3/14	
3/15		3/15	
3/16		3/16	
3/17		3/17	
3/18		3/18	
3/19		3/19	
3/20		3/20	
3/21		3/21	
3/22		3/22	
3/23		3/23	
3/24		3/24	
3/25		3/25	
3/26		3/26	
3/27		3/27	
3/28		3/28	
3/29		3/29	
3/30		3/30	
3/31		3/31	
4/1		4/1	
4/2		4/2	
4/3		4/3	
4/4		4/4	
4/5		4/5	
4/6		4/6	
4/7		4/7	
4/8		4/8	
4/9		4/9	
4/10		4/10	

Table 5: Recall rate and precision rate for cherry blossom viewing using interpolation method and optimizing of conditions (Meguro Ward)

	Recall	Precision
2018	87.5%	53.8%
2019	100.0%	41.2%

Comparison of Tables 1 and 4 reveals that the recall rate in 2018 improved from 62.5% to 87.5%. Therefore, we used a

method that combines methods of interpolation and steady state judgment to verify whether the best time to see cherry blossoms can be estimated for other wards. Figure 5 presents results obtained for cherry blossoms in Chiyoda Ward in 2018. Figure 6 presents results obtained for cherry blossoms in Sumida Ward in 2018. Table 5 shows the recall rate and the precision rate based on the best-time estimation results of cherry blossoms in Fig. 5 and Fig. 6.

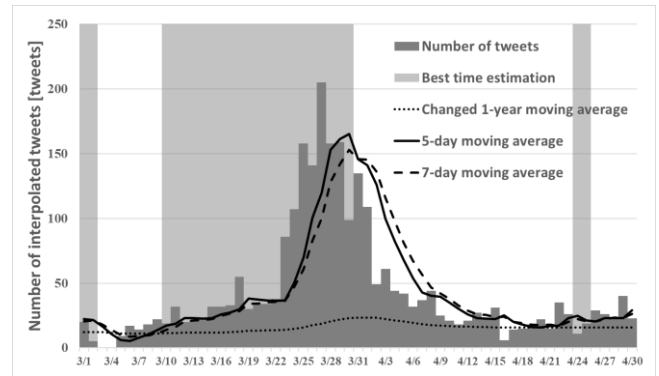


Figure 5: Chiyoda in 2018.

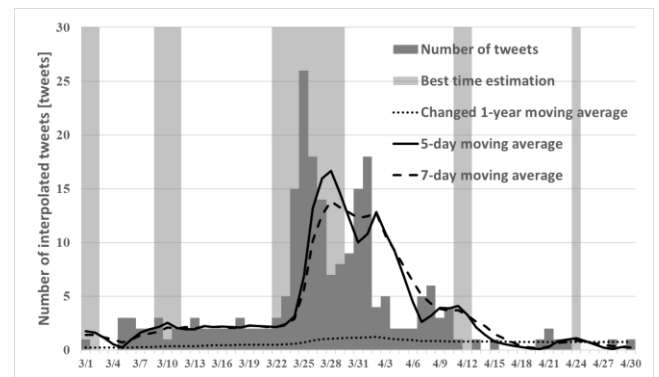


Figure 6: Sumida in 2018.

Table 6: Recall rate and precision rate for cherry blossom viewing (Chiyoda Ward and Sumida Ward in 2018)

	Recall	Precision
Chiyoda	100.0%	34.8%
Sumida	37.5%	14.3%

Table 6 shows the recall rate in Chiyoda Ward as 100%, but that in Sumida Ward as 37.5%, which was not good. The cause of the low recall rate of Sumida Ward is described in the next section.

### 5.1.1. Cause of the low recall rate of Sumida Ward

Two reasons might explain the low recall rate that occurs for Sumida Ward. The first is the small total number of tweets. The total numbers of tweets from Mar. 1 through Apr. 30 in 2018 were 1071 in Meguro Ward and 2643 in Chiyoda Ward, whereas the number was 213 in Sumida Ward: they were extremely few. Because of the small number of tweets, it is

possible that the conditions for estimating the best time did not work.

The second is the lack of parameters added. In the proposed method, only weather information was added. However, it might be necessary to consider not only weather information but also meteorological information other than the weather information, and people flow data. Because of these effects, the best time to estimate the cherry blossoms in Sumida Ward was not good.

## 6. TIME SERIES FORECASTING RESULTS

This chapter presents results of automatic prediction of the number of tweets in a certain period using a method with machine learning. First, as comparison targets, Figs. 7–9 show the best times to see cherry blossoms in Tokyo, Kyoto, and Shizuoka prefectures in 2019, as inferred using the existing method.

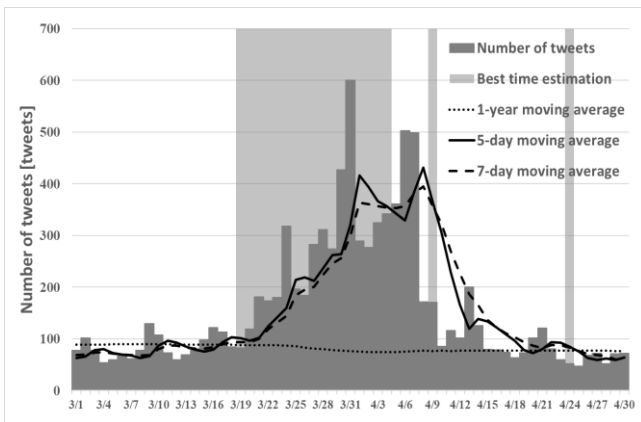


Figure 7: Tokyo in 2019.

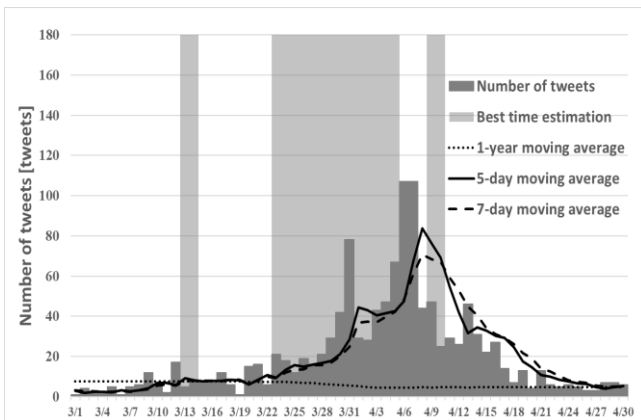


Figure 8: Kyoto in 2019.

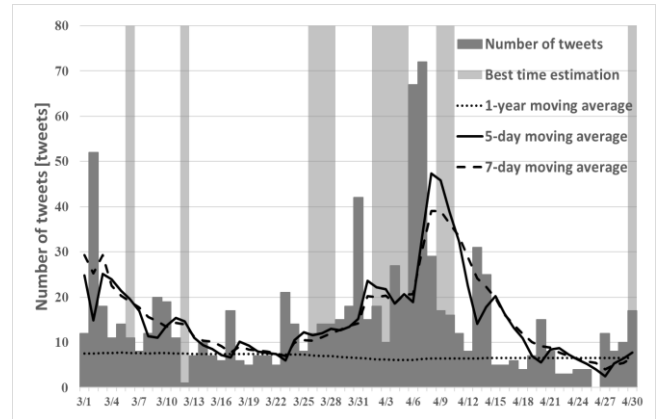


Figure 9: Shizuoka in 2019.

Next, using the existing method and the number of tweets forecasted in time series, the best time to view the cherry blossoms was estimated. Using learning data including tweet information from Feb. 1 through Apr. 30 during 2015–2018, and Feb. 1 through Feb. 28 in 2019, we predicted the number of tweets from Mar. 1 through Apr. 30, 2019. Figures 10–12 show the best times to view cherry blossoms, as obtained using the forecasted values of machine learning for Tokyo, Kyoto, and Shizuoka prefectures in 2019.

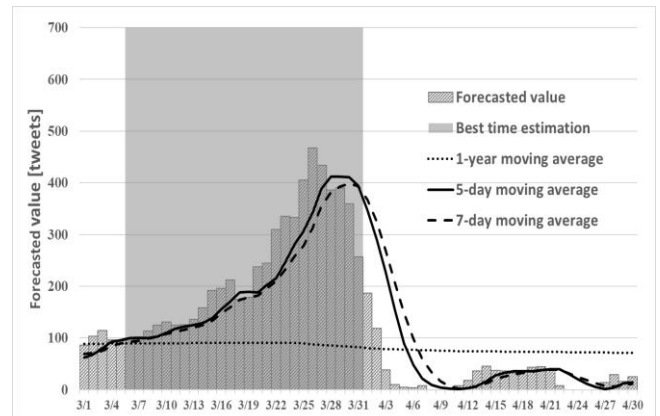


Figure 10: Tokyo after forecasting in 2019.

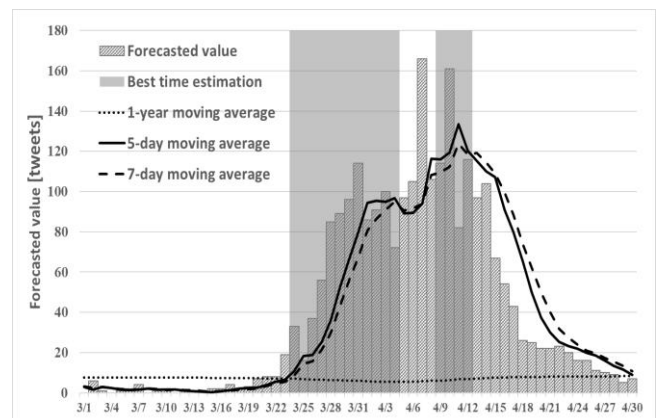


Figure 11: Kyoto after forecasting in 2019.

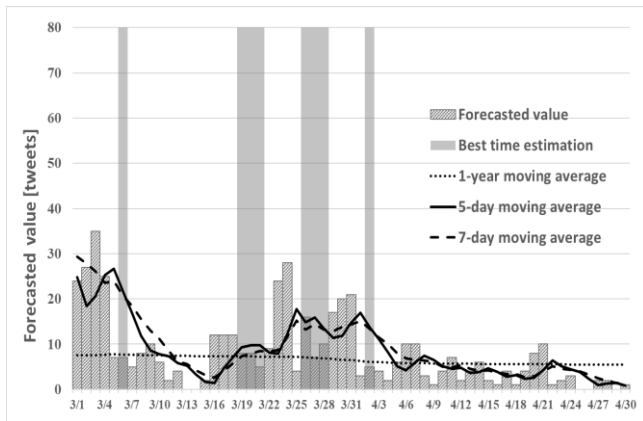


Figure 12: Shizuoka after forecasting in 2019.

Tables 7 and 8 respectively show the recall rate and precision rate of existing methods using the actual number of tweets and the predicted number of tweets.

Table 7: Recall rate and precision rate of existing methods using actual numbers of tweets

	Recall	Precision
Tokyo	100%	36.8%
Kyoto	90.9%	52.6%
Shizuoka	50.0%	29.4%

Table 8: Recall rate and precision rate of existing methods using predicted numbers of tweets

	Recall	Precision
Tokyo	100%	26.9%
Kyoto	81.8%	50.0%
Shizuoka	20.0%	12.5%

Comparison of Figs. 7–9 and Figs. 10–12 shows that the number of tweets can be predicted, although the lack of learning data and the lack of prediction locations are issues left for future studies. In particular, in the figure for Shizuoka in Fig. 12, it can be confirmed not only that the peak is in mid-March, but also that the cherry blossoms at the beginning of March are present because early blooming cherry blossoms (early February through early March) such as “Kawazu cherry blossoms” and “Kakegawa cherry blossoms” are famous in Shizuoka prefecture. In addition, from Tables 7 and 8, one can confirm that the number of tweets can be predicted, although optimization of conditions for estimating the best time in time series forecasting is a future issue. Therefore, although lack of learning data and prediction location and optimizing of conditions for estimating the best time in time series forecasting remains as an issue for future study, one can use machine learning to predict the number of tweets automatically in a certain period.

## 7. CONCLUSION

This survey was administered with the aim of improving estimation of the best time to see cherry blossoms. In earlier

studies, it became possible to estimate the best time to see cherry blossoms in each prefecture. However, when estimating the best time to see cherry blossoms at each tourist spot, a difficulty arose: the estimation accuracy was reduced significantly because of the influence of the weather. Therefore, we examined whether the accuracy of the best time to estimate the cherry blossoms at each tourist spot could be improved by adding weather information, which was a difficulty of the existing method. The best time to see the cherry blossoms in Meguro Ward, where the estimation accuracy has dropped considerably, was re-estimated using the proposed method. Results demonstrate the possibility of drawing a graph with less influence of bad weather when weather information was examined, but no change was found in the recall rate or the precision rate. Next, we specifically examined the total number of tweets for each year. The difference in the numbers of tweets in respective years was large. Therefore, we examined optimization of conditions for estimating the best time. The best time to see the cherry blossoms was estimated using the interpolation method and optimization of conditions for estimating the best time. Results show that the recall rate improved from 62.5% to 87.5% in Meguro Ward compared to the existing method. Therefore, for comparison, the best times to see cherry blossoms in Chiyoda Ward and Sumida Ward were estimated. Results show that the recall rate was 100% in Chiyoda Ward, but the recall rate was 37.5% in Sumida Ward. The cause of the low recall rate of Sumida Ward is related to the small number of tweets and the lack of parameters added.

We also investigated the possibility of using machine learning as a new method for estimating the best time to view cherry blossoms. Results confirmed the possibility of predicting the number of tweets. However, inadequacies of learning data, prediction location, and optimization of conditions for estimating the best time in time series forecasting remain as issues for future research.

For future investigations, it will be necessary to estimate the best time to view cherry blossoms in places where tweets are few. It will also be necessary to consider not only weather information but also meteorological information other than weather information, as well as people-flow data. Additionally, we will increase the learning data and prediction location and verify whether the number of tweets can be predicted by spot and by prefecture. Furthermore, at present, to secure numerous tweets, the words "さくら", "サクラ", and "桜" are targeted. However, from now on, it will be necessary to infer what states the cherry blossoms are showing from the Twitter text.

## References

- [1] Twitter, <https://twitter.com/>
- [2] Masaki Endo, Keisuke Mitomi, Keisuke Saeki, Yo Ehara, Masaharu Hirota, Shigeyoshi Ohno, and Hiroshi Ishikawa. 2016. Study of information provided by the best time to see estimation method of phenological observations using tweets. Proceedings of the 12th Annual Conference on Japan Society for Tourism Informatics. Shizuoka, Japan. pages 47–60. (in Japanese)

[3] Munenori Takahashi, Masaki Endo, Shigeyoshi Ohno, Masaharu Hirota, Hiroshi Ishikawa, 2020, Automatic detection of tourist spots and best-time estimation using social network services, International Workshop on Informatics 2020

[4] Shogo Maenaka, Shigeya Morishita, Daichi Nagata, Morihiko Tamai, Keiichi Yasumoto, Toshinobu Fukukura and Keita Sato, 2015, SakuraSensor: a system for realtime cherry-lined roads detection by in-vehicle smartphones, Proceedings of the 15th Annual Conference on the 15th Forum on International Symposium on Wearable Computers (ISWC'15). pages 345–348.

[5] Giambattista Amati, Simone Angelini, Giorgio Gambosi, Daniele Pasquin, Gianluca Rossi, Paola Vocca, 2018, Twitter: temporal events analysis: Extended Abstract, Proceedings of the Fourth EAI International Conference on Smart Objects and Technologies for Social Good. pages 298-303.

[6] Peilin Yang, Srikanth Thiagarajan, Jimmy Lin, 2018, Robust, Scalable, Real-Time Event Time Series Aggregation at Twitter, Proceedings of the 2018 International Conference on Management of Data. pages 595–599.

[7] Twitter Developers, <https://dev.twitter.com/>.

[8] Agricultural Research Institute, Simple Reverse Geocoding Service  
<https://www.finds.jp/rgeocode/index.html.ja>.

[9] goo weather, <https://weather.goo.ne.jp>

[10] AWS, <https://aws.amazon.com/jp/console/>



# Maintaining Soundness of Social Network by Understanding Fake News Dissemination and People's Belief

Risa Kusano<sup>\*,1</sup>, Kento Yoshikawa<sup>\*</sup>, Hiroyuki Sato<sup>\*</sup>, Masatsugu Ichino<sup>\*</sup>, and Hiroshi Yoshiura<sup>\*,2</sup>

<sup>\*</sup>Graduate School of Informatics and Engineering, University of Electro-Communications, Japan

<sup>\*\*</sup>Faculty of Engineering, Kyoto Tachibana University, Japan

<sup>1</sup>r1.kusano@uec.ac.jp, <sup>2</sup>yoshiura@tachibana-u.ac.jp

**Abstract** - Fake news in social networks seriously damage the society. Although main countermeasure against fake news dissemination is detection of them, previous detection methods could be circumvented once their algorithms were known to the attackers. Furthermore, they are not effective enough because users of social networks believing fake news tend to adhere to their belief and to reject opposite truth.

In this paper, we propose a detection method based on a model of simulating fake news dissemination in a social network. The Fake news Dissemination Model (FDM) is constructed by extending an opinion sharing model called Autonomous Adaptive Tuning (AAT) to capture features of fake news. The FDM has parameters such as probability of the news being true, degree of each network participant believing the news, and degree of participant believing each other. These parameters are optimized to achieve the matching state where the FDM accords with dissemination data of news to be checked. Based on the values of the parameters in the matching state, the proposed method determines whether the news is real or fake as well as estimates how strongly each participant believe the news and other participants, from which we can plan how to persuade the participants to accept the truth.

We have implemented a simple version of the proposed method. Preliminary experiments using real data of fake news dissemination revealed features of the proposed method and future work to make the method practical.

**Keywords:** Fake news, Detection, Correction, Persuasion, Dissemination model

## 1 INTRODUCTION

Means by which people receive and disseminate information have changed from mass media (e.g. newspapers, television) to social media (e.g. Facebook, Twitter). Social media enable people to disseminate information rapidly and without formal gatekeepers. As a result, deceptive and misleading news, i.e. fake news, has become widespread, which negatively affects individuals and society [1].

A representative countermeasure against fake news is fake news detection by analysing the texts and images used for the news items [2], checking the speed at which the news spreads [3], and evaluating the reliability of the people reporting the news [4]. However, widespread communication on social media has amplified the echo chamber effect in which one's beliefs are strengthened through interactions with like-minded individuals [5]. It has also amplified the backfire effect in which facts that contradict one's beliefs are rejected

[5]. Therefore, simple detection is not sufficient to prevent dissemination of fake news because people who believe fake news tend to adhere to fake news irrespective of the truth.

In this paper, we present a countermeasure against fake news that not only detects fake news items but also estimates the degree of each person believing the news items and of each person believing the people who send the news items to him or her. The detection and estimation are done in a single framework based on matching between the real data of news item dissemination and a model of fake news dissemination. Our contributions are summarized as follows.

We propose a novel countermeasure framework targeting fake news that is based on matching between news item dissemination data and a fake news dissemination model (FDM). It should enable not only detection of fake news items but also estimation of the degree of each person believing the news items and believing their friends. This will enable identification of the people who can be persuaded to change their belief about fake news and identification of the people who are the best persuaders. The framework also enables prediction of further dissemination of fake news.

## 2 RELATED WORK

### 2.1 Fake news detection

A representative countermeasure against fake news is its detection. Detection methods can be classified into three types.

Text and images used for fake news items have features different from those used for real news items. For example, text used for fake news typically includes hearsay (e.g. "I heard from a friend that...") or speculation (e.g. "maybe") [6]. Detection methods of the first type focus on the features of social media posts mentioning news [2, 7]. However, these methods can be circumvented by avoiding expressions and images typical of fake news items.

Fake news is known to spread in a network differently from real news. For example, fake news tends to spread faster than real news. Detection methods of the second type focus on how it has spread in a social network [3, 8]. However, these methods can also be circumvented by tailoring posts mentioning fake news to spread like ones for real news by, for example, making the wording less inflammatory and

having them posted by several colluders<sup>1</sup>. Another problem with these methods is that they work only after the news has been spread in a network.

The reliability of a person does not drastically change in the short term, so posters can be divided into those who mostly post or forward real news items and those who mostly post or forward fake news items. Detection methods of the third type are based on the relationship between people and news [4, 9]. These methods take a set of news items and a set of persons and determine the reliability of each item and that of each person to maximize a likelihood function. Although these methods are more difficult to circumvent than those of the first two types, they work only after the news has been spread in a network.

All three types of detection methods share a common problem—people who believe fake news tend to adhere to fake news irrespective of the truth. We therefore present a new type of detection method that is based on matching between data of news item dissemination and our model of fake news dissemination, i.e. the FDM.

## 2.2 Opinion sharing model

Because our FDM is an extension of an opinion sharing model, we survey opinion sharing models. Opinion sharing models have been used to explain how users in social networks share opinions. These models can be classified into two types: those that assume the presence of the ground truth (i.e. a fact that supports or contradicts an opinion) [10, 11] and those that do not [12, 13, 14].

The former type describes the dissemination of opinions that match (or mismatch) the ground truth. Ginton et al. modelled the “opinion sharing problem” in which users share opinions through local interaction among users [10]. Prymak et al. improved the precision of Ginton’s model by developing an opinion dissemination model based on an autonomous adaptive tuning (AAT) algorithm [11].

The latter type does not consider whether an opinion matches the ground truth and simply describe how people change their opinions due to communication. DeGroot proposed a model in which each user updates his or her opinions by weighted average of his/her current opinion and opinions of his/her friends in accordance with the importance of their friends [12]. Tsang and Larson extended this model to explain opinion transition in which diverse opinions converge to a few representative opinions by adding people who never change their opinions [14].

## 2.3 Fake news dissemination model

Of the two types of opinion sharing models, the ones in which the ground truth is assumed are suitable as a basis for the FDM because we can determine whether news is real or fake by comparing the news with the ground truth. We used Prymak’s model based on AAT as a representative example of this type because the AAT algorithm is more accurate and faster than the algorithms used by other models of this type.

The AAT algorithm uses three parameters: the subjective probability of each user  $i$  believing that the news item is true ( $P_i$ ), the degree to which each user  $i$  believes his/her friend ( $t_i$ ), and the degree to which each user  $i$  doubts his/her friend ( $f_i$ ), where  $1 \leq i \leq M$ , and  $M$  is the number of users. Prymak et al. modelled the dissemination of correct and incorrect news but does not consider fake news that is not only incorrect but also fabricated with malicious intent such as changing election results. We have extended Prymak’s model as follows [15].

- In Prymak’s model, it is assumed that each user  $i$  believes or doubts all his/her friends equally. This assumption is not valid when one friend sends a real news item and another friend sends a fake news item. We therefore use the degree to which each user  $i$  believes each of his/her friends and use  $t_{ij}$  instead of  $t_i$ . Similarly, we use  $f_{ij}$  instead of  $f_i$ .

- In Prymak’s model, it is assumed that  $t_i$  and  $f_i$  are constant during opinion dissemination in a network. This assumption is not valid when users recognize fake news. We therefore use a mechanism, the expectation maximization (EM) algorithm, to updated  $t_{ij}$  and  $f_{ij}$  during news dissemination.

We overview FDM by using Fig 1. Each user has a belief about the news (i.e. the subjective probability of the user believing that the news is real), a belief in the reliability of each of his/her friends, and a doubt about each of his/her friends at each status of the social network. Each user receives an opinion about a news item from one or more of his/her friends. He/She updates the reliability and doubt for each friend on the basis of these opinions. The user then updates his/her belief about the news by using the Bayes update formula on the basis of each opinion received, the corresponding updated reliability, and the corresponding updated doubt (see [15] for details).

For example, in Fig. 1, at status  $k$ , user 4 receives opinions about a news item from users 1 and 2. Both opinions are that the news item is real. User 4 believes user 1 with a degree of 0.45 ( $t_{41}^{k-1}$ ) and doubts him/her with a degree of 0.40 ( $f_{41}^{k-1}$ ). User 4 similarly believes and doubts user 2. In accordance with the opinions received and the reliability and doubt for each friend who sent an opinion, user 4 updates the reliability and doubt for users 1 and 2 from  $(t_{41}^{k-1}, f_{41}^{k-1}, t_{42}^{k-1}, f_{42}^{k-1})$  to  $(t_{41}^k, f_{41}^k, t_{42}^k, f_{42}^k)$ . He/She changes his/her belief about the news item from  $P_4^k=0.70$  to  $P_4^{k+1}=0.85$  on the basis of the opinions, the updated reliability, and the updated doubt. Because user 4’s belief about the news item exceeds the threshold,  $\sigma=0.80$ , he/she takes the opinion that the news item is real | true and shares it with users 5 and 7.

The formula for the Bayes update is symmetric, i.e.  $P_i$  has a value  $p$  with parameters  $t_i$  and  $f_i$  if and only if  $1 - P_i$  has the same value  $p$  with parameters  $f_i$  and  $t_i$ .

In this paper, we describe how to use the FDM to detect fake news, persuade people to disbelieve it, and predict its future dissemination.

<sup>1</sup>Weak attacks by many colluders are actually used for network intrusion and thus could be adapted for fake news dissemination.

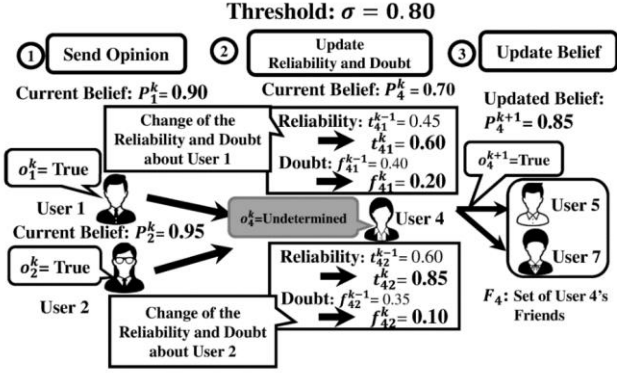


Fig. 1. Overview of fake news dissemination model [15]

### 3 PROPOSED METHOD

#### 3.1 Assumption

Our method depends on two assumptions.

- 1) Our FDM is correct enough; i.e. the nearer the simulated news dissemination data in FDM (with a vector  $V$  of its parameter values) to the real news dissemination data, the more precisely vector  $V$  represents the network situation.
- 2) The parameters in  $V$  can be optimized in reasonable time so that the simulated news dissemination data in FDM are near enough to the real news dissemination data.

#### 3.2 Method

We call the real news dissemination data “real data” and the news dissemination data simulated using FDM and its parameter vector  $V$  “simulated data”. The quantified difference between the real and simulated data is the evaluation function  $F(V)$  and the smaller  $F(V)$ , the more precisely the parameters in  $V$  represents the network situation.

Figure 2 illustrates the flow of our method. First, vector  $V$  of the values of the FDM parameters is initialized. Then, the simulated data are calculated using FDM and  $V$ , and the evaluation value  $F(V)$  is calculated. The parameters in  $V$  are repeatedly updated to minimize  $F(V)$  until  $F(V)$  is less than or equal to the threshold value  $T1$ .

## 4 IMPLEMENTATION

#### 4.1 Data and their representation

We used dissemination data for  $N$  news items. These data included the content of each news item, the number  $M$  of users in a social network who sent (originated or forwarded) at least one news item, the links between users, and the relationships between the users and each news items (sent, ignored, or did not see).

These data are represented by two tables. The one shown in Fig. 3 represents the relationships between the users and news items (rows represent users; columns represent news items). Each  $(i, j)$  element in the table has a 1, -1, or 0, which

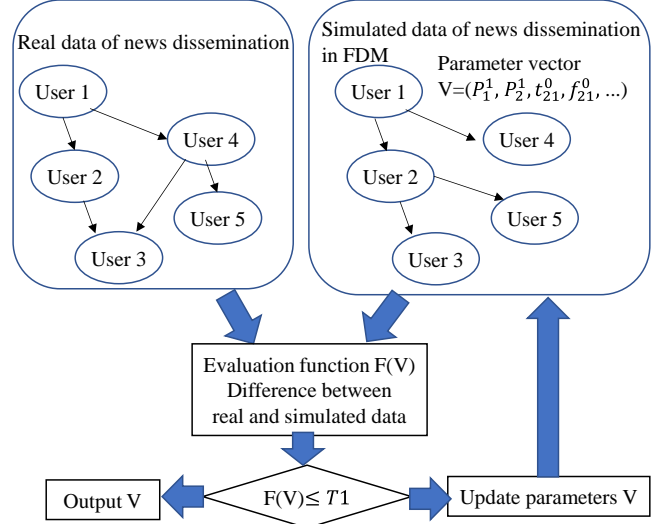


Fig. 2. Overview of proposed method

means that user  $i$  sent, ignored, or did not see news item  $j$ . The other table (not shown) represents the links between users, where element  $(i, k)$  has a 1 or 0, which means that user  $k$  follows or does not follow user  $i$  (the link is directional). A bidirectional link between users  $i$  and  $k$  (e.g. a friend link in the Facebook network) is represented by two elements  $(i, k)$  and  $(k, i)$  both having 1.

	1	2	...	j	...	N
1	1	0	...	-1	...	0
2	0	-1	...	0	...	1
...	...	...	...	...	...	...
i	-1	0	...	1	...	0
...	...	...	...	...	...	...
M	0	0	...	0	...	-1

Fig. 3. Relationship between users and news items

#### 4.2 FDM simulation

We generated an FDM with the same number of users ( $M$ ) and the same links between the users as those in the real data. Thus, the generated FDM had the same graph structure as that of the real data. We then initialized parameter vector  $V$  of the FDM. From the real data, we identified the  $M'$  users who sent the  $N$  news items and had the corresponding  $M'$  users in the FDM send the  $N$  news items (the set of these  $M'$  users are a subset of all  $M$  users). The dissemination of the  $N$  news items was simulated in the FDM; i.e. each user who received a news item forwarded it or ignored it in accordance with the parameter values. The order of forwarding the items was recorded in the real data and the same order was followed in the simulation.

Table 1: News items used for preliminary evaluation

Fake 1	<Fact Check Initiative> Dr. *** said PCR test should not be used for diagnosing infectious disease (2020.08.**)
Fake 2	<Fact Check Initiative> The WHO confirmed that Covid-19 infections have ended in Japan. (2020.08.**)
Real 1	<Breaking news> Comedian *** infected with Covid-19. He was positive at PCR test. His office disclosed on its Web page. (2020.8.**)
Real 2	Athlete *** tested positive with Covid-19 last weekend. Jamaican Ministry of Public Health disclosed. (2020.08.**)

### 4.3 Parameter optimization

We used differential evolution, a branch of evolutionary computation, to implement parameter optimization to minimize  $F(V)$  [16]. We used the DEAP software library [17] for the differential evolution.

### 4.4 Detection algorithm

The detection algorithm comprises six steps. The stop condition in Step 5 is either  $F(V)$  is smaller than threshold T1 or Steps 3 through 6 have been repeated more than threshold T2.

- Step 1: Set up FDM with the same structure as that of the real network
- Step 2: Initialize  $V$  (FDM parameters)
- Step 3: Simulate dissemination of news items
- Step 4: Calculate difference  $F(V)$  between simulated and real disseminations
- Step 5: Stop and output  $V$  if a stop condition is met
- Step 6: Update parameters by using differential evolution and go to Step 3

### 4.5 Use of parameters for persuasion

The values in parameter vector  $V$  at the point of fake news detection (Step 5 in Section 4.4) are most likely values that represent the degree of each user believing the news items and the degree of each user believing and doubting his/her friends. It should be possible to use these parameter values to persuade users to accept the truth.

## 5 PRELIMINARY EVALUATION

### 5.1 Data set

The data set used in our preliminary evaluation included two real news items and two fake news items taken from Twitter (original tweets were in Japanese). They are shown in Table 1. The two fake news items were found on the Fact Check Initiative website [18]. Approximately 11,900 Twitter users sent and/or received the four news items. The number of Follow-Follower links between the users was 126,137.

<sup>2</sup>Our algorithm outputs two symmetric interpretations due to the symmetricity of the Bayes update formula described in Section 2.3.

### 5.2 FDM and algorithm

For our preliminary evaluation, we simplified the FDM so that each user believes and doubts all of his/her friends equally, i.e. we used  $t_i$  and  $f_i$  instead of  $t_{ij}$  and  $f_{ij}$ , where indexes  $i$  and  $j$  represent the  $i$ -th user and his/her  $j$ -th friend.

We first assigned Real or Fake to each of the four news items using four-digit representation (0011, 0101, etc.). There were thus 16 assignment possibilities. For each one, we executed the algorithm shown in Section 4.4. We used  $F(A, V)$  to represent  $F(V)$  with assignment of Real or Fake  $A$ . The assignment  $A$  and parameter vector  $V$  that gave the lowest evaluation value  $F(A, V)$  was taken as the detection result. This was done using the following expression (1),

$$\text{Argmini}_{A \in \text{All\_assignments}, V \in \text{All\_parameter\_vectors}} F(A, V) \quad (1)$$

where  $\text{All\_assignments}$  and  $\text{All\_parameter\_vectors}$  respectively represent the set of 16 assignments of Real or Fake to the 4 news items and the set of all possible parameter vectors.

### 5.3 Results

Table 2 shows the detection results. The two correct solutions, i.e. FFTT and TTFF, were ranked first and fourth<sup>2</sup>. Although the two solutions should theoretically have the same evaluation value  $F(V)$ , the values differed. This is attributed to the values not having sufficiently converged.

Figure 4 shows the distributions of degrees of each user believing the four news items. The horizontal axes show the degree of a user believing the news item, and the vertical axes show the number of users who had the corresponding belief degree. Most users were neutral about believing the news while some strongly believed or rejected the news. For the two fake news items, the number of people who strongly rejected them is larger than the number of people who strongly believed them. For the two real news items, the numbers are on the contrary.

Users who weakly believed the fake news can be identified based on the degree of their believing the news. For example, we can identify 312 users whose degrees of believing the fake

Table 2: Detection results

Estimated truth of four news items	Evaluation value $F(A, V)$
TFF	0.29215
FTFF	0.29792
TTFT	0.31578
FFTT	0.31707
TFTT	0.32768
FTFT	0.33658
TFTF	0.33859
FTTF	0.34160
FFTF	0.35293
FFFF	0.41549
TTTT	0.42900
FTTT	0.43341
FFFT	0.43761
TFFT	0.45174
TFFF	0.45748
TTTF	0.46442

news 2 were in the range of (0.525, 0.625). We may be able to persuade those users to disbelieve the fake news 2.

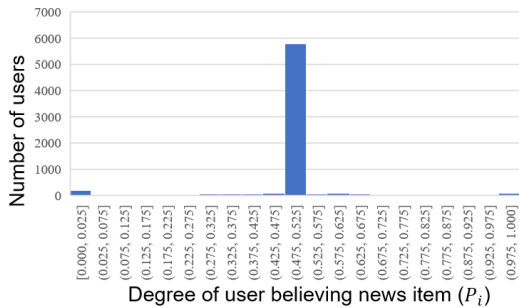
Figures 5 (a) and (b) show the distribution of degree of each user believing and doubting his/her friends. Target users for persuasion are those who weakly believe the fake news and believe their friends. Those friends can be asked to tell the truth to the target users. Other target users are users who

weakly believe the fake news and doubt their friends. These users can be told the truth through mass media.

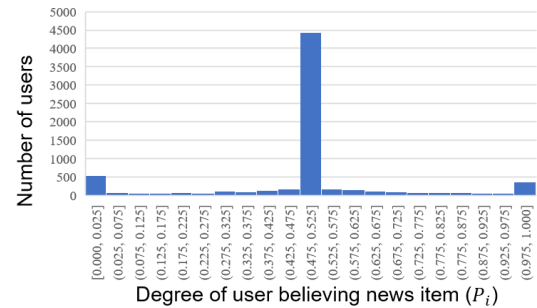
## 6 DISCUSSION

If elements with a value 1 or  $-1$  in the  $j$ -th column of Table 1, which represents news item  $j$ , are disjoint with those in the  $k$ -th column (news item  $k$ ), news items  $j$  and  $k$  disseminated in the two disjoint user groups. Therefore, the truths of news items  $j$  and  $k$  are independent. If all  $N$  news items are mutually independent in this way, their truths are mutually independent. Our algorithm depends on only the relationships between users and news items and not on the features of each news item. Thus, our algorithm arbitrarily assigns True or Fake to each of the  $N$  news items that are mutually independent.

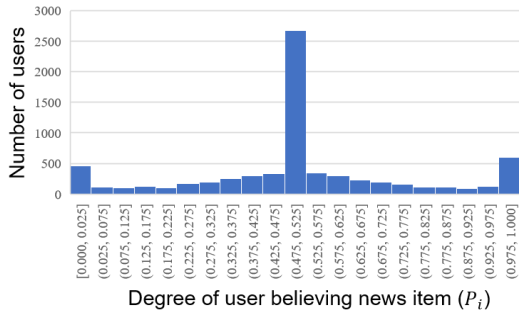
Thus, our algorithm is valid only when the news items are not independent, i.e. when elements with a value 1 or  $-1$  in a column in Table 1 overlap with those in at least one other column. Assume that elements  $(i, j)$  and  $(i, k)$  have a common value, 1; i.e. user  $i$  sent both news items  $j$  and  $k$ , meaning that user  $i$  believes both news items, which came from one friend or two different friends. Because user  $i$  believed or doubted all of his/her friends equally in our evaluation, our algorithm enables only two interpretations; i.e. both news items are real and user  $i$  believes one or both friends (Fig. 6 (a)), or both news items are fake and user  $k$  doubts one or both friends (Fig. 6 (b)). Similarly, if element  $(i, j)$  has value 1 and  $(i, k)$  has 0,



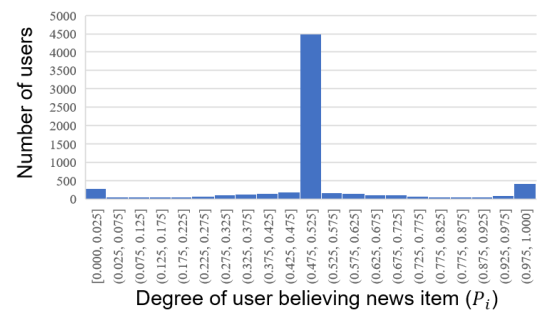
(a) Fake 1



(b) Fake 2



(c) Real 1



(d) Real 2

Fig. 4. Degrees of users believing four news items

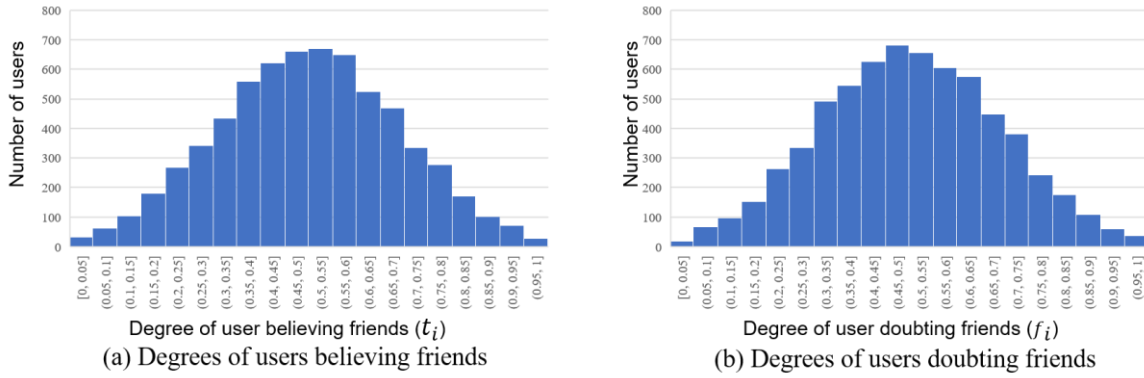


Fig. 5. Degrees of users believing and doubting friends

our algorithm estimates either that news item  $j$  is real, news item  $k$  is fake, and user  $i$  believes all friends, or that news item  $j$  is fake, news item  $k$  is real, and user  $k$  doubts all friends.

As mentioned in Section 2.3, a user seems to believe each of his/her friends differently instead of believing them equally. If so, our algorithm could arbitrarily assign True or Fake to each news item. Assume, for example, that user  $i$  believes two news items that came from different friends. If the user believes or doubts the two friends differently, our algorithm enables all four interpretations as follows and thus can determine nothing.

- 1) Both items are real, and user  $k$  believes all friends (Fig. 6 (a)).
- 2) Both items are fake, and user  $k$  doubts all friends (Fig. 6 (b)).
- 3) News item  $j$  is real, news item  $k$  is fake, user  $i$  believes the friend who sent news item  $j$ , and he/she doubts the friend who sent news item  $k$  (Fig. 7 (a)).
- 4) News item  $j$  is fake, news item  $k$  is real, user  $i$  doubts the friend who sent news item  $j$ , and he/she believes the friend who sent news item  $k$  (Fig. 7 (b)).

To avoid this problem, our algorithm should not let a user believe his/her friends fully independently but rather should incorporate adequate probabilistic distributions of the degree of a user believing his/her friends. Furthermore, it may be

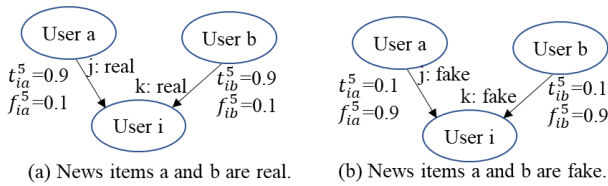


Fig. 6. Detection example with uniform reliability and doubt

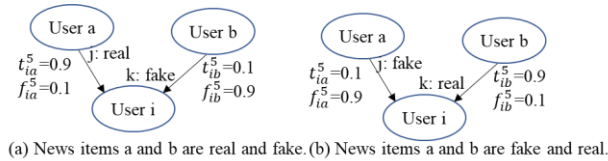


Fig. 7. Detection example with different reliability and doubt

necessary to train our FDM before use so that the degrees of users believing their friends have realistic initial values.

## 7 CONCLUSION

Conventional countermeasure against fake news simply detect it and thus fail to lead people to accept the truth. Our novel framework of countermeasures enables fake news not only to be detected but also the degree of network users believing the news and of believing their friends to be estimated. This will facilitate identification of the users who can be persuaded to accept truth and identification of the people who are the best persuaders. The framework also enables prediction of further dissemination of the fake news.

The proposed framework optimizes parameter values so that the simulated news dissemination data are nearest to the real data. If the optimized parameter values represent the situation of the real network in which the news is disseminated, the probability of a news item being true, the degree of users believing the news item, and the degree of users believing his/her friends can be determined.

A simplified version of our method was implemented, and preliminary evaluation with four news items (including two fake new and real news items) demonstrated the viability of our framework and revealed some of its problems.

## REFERENCES

- [1] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, Fake news detection on social media: A data mining perspective, *ACM SIGKDD Explorations Newsletter* 19(1), 22–36 (2017).
- [2] Y. Wang, et al., EANN: Event adversarial neural networks for multi-modal fake news detection, In: *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 849–857 (2018).
- [3] S. Vosoughi, D. Roy, and S. Aral, The spread of true and false news online, *Science* 359(6380), 1146–1151 (2018).
- [4] S. Yang, et al., Unsupervised fake news detection on social media: A generative approach, In: *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 1, pp. 5644–5651 (2019).



- [5] W. Quattrociocchi, A. Scala, C.R. Sunstein, Echo chambers on Facebook, Available at SSRN 2795110 (2016).
- [6] S. Kwon, M. Cha, and K. Jung, Rumor detection over varying time windows, *PloS one* 12(1), e0168344, (2017).
- [7] A. Gupta, H. Lamba, P. Kumaraguru, and A. Joshi, Faking Sandy: characterizing and identifying fake images on twitter during Hurricane Sandy, In: *Proceedings of the 22nd International Conference on World Wide Web*, pp. 729–736 (2013).
- [8] X. Zhou, R. Zafarani, Network-based Fake News Detection: A Pattern-driven Approach, *ACM SIGKDD Explorations Newsletter* 21(2), 48–60 (2019).
- [9] N. Sitaula, C.K. Mohan, J. Grygiel, X. Zhou, and R. Zafarani, Credibility-based Fake News Detection, In *Disinformation, Misinformation and Fake News in Social Media: Emerging Research Challenges and Opportunities*. Springer, 163–182 (2020).
- [10] R.T. Glinton, P. Scerri, and K. Sycara, Towards the understanding of information dynamics in large scale networked systems, In: *Proceedings of the 12th International Conference on Information Fusion*, IEEE, pp. 794–801 (2009).
- [11] O. Prymak, A. Rogers, and N.R. Jennings, Efficient opinion sharing in large decentralised teams, In: *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems*, pp. 543–550 (2012).
- [12] M.H. DeGroot, Reaching a consensus, *Journal of the American Statistical Association* 69(345), 118–121 (1974).
- [13] K. Sasahara, et al., Social influence and unfollowing accelerate the emergence of echo chambers, *Journal of Computational Social Science* 4, pp. 381–402 (2020).
- [14] A. Tsang, K. Larson, Opinion dynamics of skeptical agents, In: *Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems*, pp. 277–284 (2014).
- [15] K. Yoshikawa, T. Awa, R. Kusano, H. Sato, M. Ichino, and H. Yoshiura, A Fake News Dissemination Model Based on Updating Reliability and Doubt among Individuals, In: *Proceedings of the 11th International Conference on Awareness Science and Technology (iCAST)*, pp. 1–8 (2020).
- [16] F-A. Fortin, F-M. De Rainville, M-A. Gardner, M. Parizeau, and C. Gagné, DEAP: Evolutionary Algorithms Made Easy, *Journal of Machine Learning Research* 13, pp. 2171–2175 (2012).
- [17] DEAP, <https://github.com/DEAP/deap>
- [18] Fact Check Initiative, <https://fij.info/coronavirus-feature>

## ACKNOWLEDGEMENT

This work was supported by JSPS KAKENHI Grant Number JP21K11883.





Session 3:  
IoT and ITS  
( Chair: Yuichi Tokunaga )



# Multi-channel Communication of LoRa using Time Division Multiple Access

Sakauchi Ryotaro\*, Shuto Ishikawa\*, Hikaru Yabe\* and Mikiko Sode Tanaka\*\*

\* Kanazawa Institute of Technology, Japan

\*\* International College of Technology, Japan

{b1905429, b1816880, b1800744}@planet.kanazawa-it.ac.jp

sode@neptune.kanazawa-it.ac.jp

**Abstract** – We are working on the development of a bus operation management system using LoRa. The main part of the bus operation management system is a bus location system called the "BusDoko System" that presents the current location of the bus. In the system, a GPS and an IoT sim have been used to collect real-time location information from running buses. We are currently working on a change to LoRa communications in order to reduce the operation cost. In recent years, there are many users in the 920MHz band, and stable system operation has become difficult. Therefore, in this paper, we have examined a method that enables stable communications even under such circumstances. In order to improve the accuracy of the system, we have created a multi-channel LoRa communication method using Time Division Multiple Access (TDMA) for the bus location system. We have confirmed the effect from results of experiments in a room, so we will explain the details of the results. And we report on the successful operation of the communication system and its effectiveness against communication interference.

**Keywords:** LPWA, LoRa, Time Division Multiple Access, Channel Hopping, Bus location

## 1 INTRODUCTION

Community buses play an important role in the transportation of local residents. However, the operation of it is depend on weather and traffic conditions. Therefore, it may give the bus users anxiety about whether the bus will arrive. In recent years, efforts have been made to develop and operate the bus location systems that present the current location of buses in order to remove unstable elements [1][2][3]. Figure 1 shows the overall bus location system which we developed, the "BusDoko System". This system presents the current location of community buses Notty circling Nonoichi City, Ishikawa Prefecture, on a website.

In the system, GPS, IoT sim and LoRa module has been used to collect the real-time bus location information from running buses. In recent years, there are many users in the 920MHz band, and stable system operation has become difficult. In the system interference is a problem that needs to be resolved to improve system stability. Many wireless standards in recent years have capable of long-distance transmission, and since the timing and transmission channel can be freely set, they are susceptible to interference.

Therefore, in order to ensure the communication quality of the system, the time occupancy rate of other system signals arriving at the Gateway (hereinafter referred to as the out-of-system interference time ratio) is estimated each channel with high accuracy, and the channel of LoRa of the our system is decided based on the estimation result. We need to assign the communication between bus and gateway to the appropriate channel to maintain the system stability.

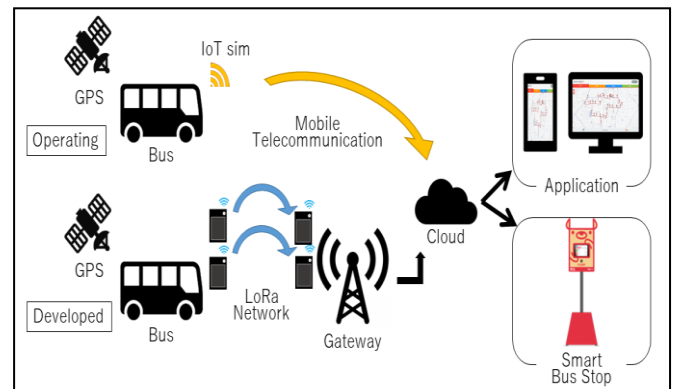


Figure 1: Bus location system which is called BusDoko system

There are existing studies [4] [5] [6] as a highly accurate radio wave environment prediction method. The method of acquiring, accumulating, and utilizing the radio wave condition is an important method from the viewpoint of radio wave utilization efficiency. In the existing studies [5] [6], the actual observation type database is discussed, and its usefulness is reported. This database is a collection of environmental information such as received signal power and reception position information acquired by the receiver in the actual environment. Using this, the communication parameters of wireless channels, and channels can be adaptively assigned. In recent years, the use of this database has been discussed in the field of frequency sharing. On the other hand, research is being conducted on LoRa to suppress interference [7] [8]. However, these do not allow the system to consider the effects of interference.

In order to improve the accuracy and reliability of the system, we have created a method for multi-channel LoRa communication method using Time Division Multiple Access (TDMA) for the bus location system. We also create the channel hopping method which use the observation type

wireless database. This makes it possible to handle changes in the communication environment over time.

We have confirmed the effect as a result of experiments in the room. We will explain the details of the experiments. And we report on the successful operation of the communication system and its effectiveness against communication interference.

## 2 BUSDOKO SYSTEM

In this section, we will explain the configuration of the BusDoko system. The BusDoko system is a system that allows bus users to check the location information of buses in the real time on a Web page, a mobile terminal, or a terminal installed at bus stops.

By using the bus location system, it is possible to visualize the arrival, the departure, and the location of buses, and it is possible to eliminate the anxiety and dissatisfaction that bus users have. Figure 1 shows the configuration of the bus location system we have developed. The buses are equipped with an in-vehicle device equipped with Private LoRa, IoT sim and GPS, acquires the current position information from the GPS at intervals of several seconds, sends it to the gateway using Private LoRa. And then gateway send data to cloud via the Internet and store the data in the upper server. By displaying the bus position information on a Web page using the stored the bus position information data, the bus position can be confirmed on the user's terminal. In addition, some of the bus stops we have developed have a built-in tablet, and some of the bus stop are equipped with a panel that allows you to check the bus position using LEDs. The LoRa gateway send the bus locations to the bus stops. Therefore the bus stops can display the current position of the buses.

Antenna power	CH number	CH used in a bundle	Carrier sense time	Sending duration	Pause duration	The sum of emission time per arbitrary 1 hour
250mW (24dBm) or less	24-32	1-5ch	5ms or more	4s	50ms	None
	33-38					360s or less
	33-38	1ch	128 $\mu$ s or more	More than 200ms, and 400 ms or less	Ten times or more of the former transmitting time or 2ms	360s or less

Figure 2: 920MHz band channel allocation specified by ARIB [9]

The domestic standard for the use of the 920MHz band is specified in ARIB STD-T 106-108. Figure 2 shows an excerpt of the ARIB STD-T108 standard. LoRa is a low-power wireless system that uses the 920MHz band, and it is possible to perform communication according to the purpose by changing the output power to 20mW or less and 250mW or less. The standard defines a combination of usable channel bandwidth, carrier sense time, maximum

transmission time, pause time, and total transmission time per hour for each of 20 mW or less and 250 mW or less. We are using 250mW LoRa in the BusDoko system.

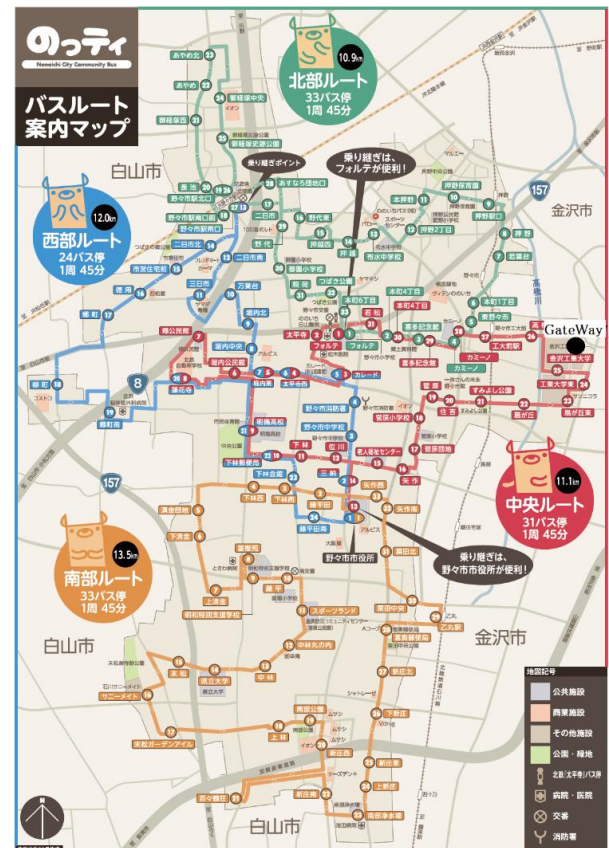


Figure 3: Bus route of the community bus Notty



Figure 4: Gateway on the Library center

Figure 3 shows the bus route map of the community bus Notty. The gateway was installed on the roof of the 12-story Kanazawa Institute of Technology Library Center. There are four bus routes for the community bus. Figure 4 shows the gateway installed on the roof of the library center. There are two boxes with four LoRa modules.

### 3 MULTI-CHANNEL COMMUNICATION

If an end device on a bus sends data using one channel, it cannot be sent if that channel is used by another system. To avoid this danger, we decided to use two channels (see Figure 5). This allows the location of the bus to be sent to the gateway, even if either channel is used by the other system.

On the other hand, due to the limited number of channels, it is not good for one end device to occupy two channels. Therefore, we decided to use two Channels in common by two end devices.

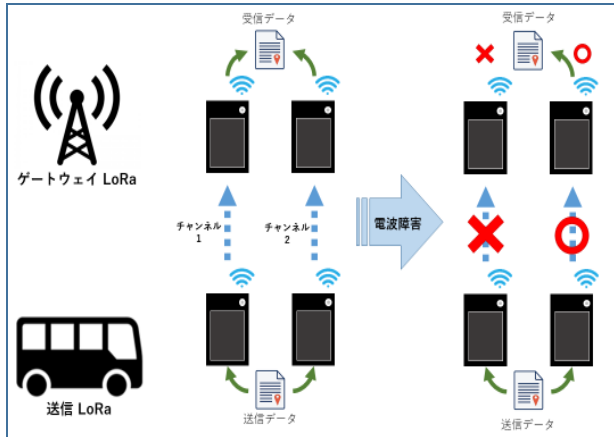


Figure 5. Effect of using multiple channels

#### A. Multi-channel Communication

In the case of communication using a single channel, regular communication may become difficult due to interference by others, which may impair the quality and real-time performance of the system. Therefore, we decided the multiple devices will communicate using different channels to solve these problems. Figure 6 shows the system diagram of LoRa communication using multi-channel in Nonoichi Notty bus routes. In this communication, two LoRa modules are installed in the end device of each bus, and the location information acquired by GPS is transmitted alternately. Each LoRa parameter sets a different channel. At the LoRa gateway, a module adapted to the channel receives the data and uses the ID assigned to the data to determine the route.

#### B. Time-division multiple access of LoRa Link

When multiple communications are performed in the same channel, data loss due to communication collision is considered. We decided to synchronize the time between buses and gateway by using the time that is gotten at the same time by acquiring the position of the bus using the GPS. By synchronizing the time in all LoRa of the end device, setting the time frame, and controlling the transmission module and time, the communication within the same channel is possible between two devices. Figure 7 shows the time frame diagram of the TDMA-based multi-ch communication. The LoRa gateway is always active and waits for transmission from the LoRa end device. The end

device synchronizes the time based on the time data obtained from GPS. The time frame starts at exactly even minutes, and the end device communicates every 4 seconds on the selected channel. By limiting the number of transmitting modules in the same time frame, one connection is established for each channel to realize TDMA.

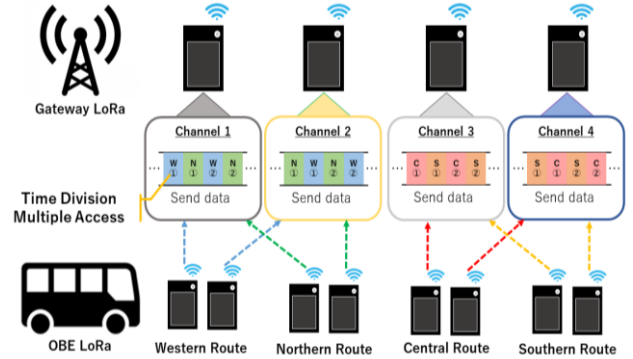


Figure 6. Channel assignment of Nonoichi nottey bus route

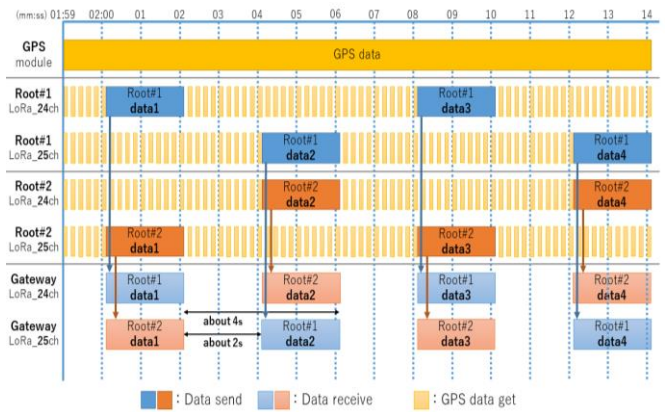


Figure 7. Multi-channel communication image

#### C. Channel hopping

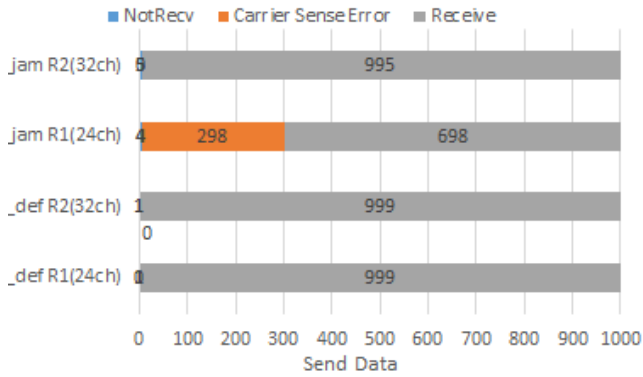
The channel used by the end device is determined using an actual observation database. The actual observation database is collected and created using a gateway data. The channel number will be updated using the time when all buses gather at Nonoichi City Hall for transfer on each bus route and stop for about 5 minutes. That is, the channel number to be transmitted is changed once an hour. This makes it possible to handle temporal channel congestion.

## 4 EXPERIMENTAL RESULTS

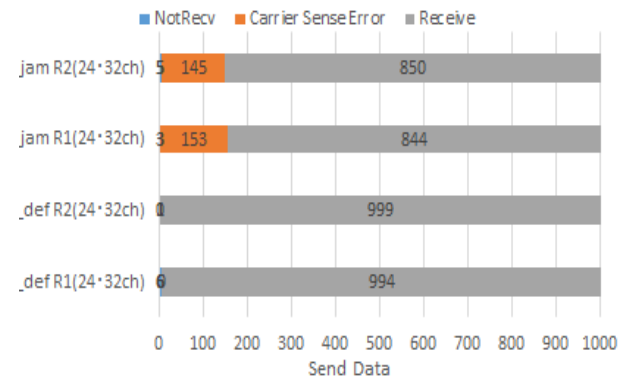
An experiment was conducted to verify the effectiveness of multi-channel communication when jamming occurs. Test data was transmitted from the end device for two routes using 24 and 32 channels, and different data was transmitted at intervals of 4~8 seconds in 24 channels of the jamming device.



Figure 8 shows the number of successful and unsuccessful transmissions when interference is given. Figure 8 (a) shows the result of an example using only one channel in the end device. Figure 8 (b) shows the results of an example using two channels in the end device. In the 1-channel communication, the number of carrier sense errors is biased toward route 1. In the 2-channel communication, the number of errors is successfully distributed.



(a) Experimental results of pattern 1  
(With / Without interference)



(b) Experimental results of pattern 2  
(With / Without interference)

Figure 8. Transmission success/failure results

From the experiments, we checked whether data could be transmitted at an acceptable time, even in situations where other systems used the same channel as the bus-end device and had collisions. From the results, it was found that the data could be transmitted within a reasonable time and the system worked.

## 5 CONCLUSION

We are working on the development of a bus operation management system using LoRa. In recent years, there are many users in the 920MHz band, and stable system operation has become difficult. Therefore, in this paper, we have examined a method that enables stable communication even under such circumstances. In order to improve the accuracy of the system, we have created a method for multi-channel LoRa communication using TDMA for the bus

location system. We also created the channel hopping method. This makes it possible to handle changes in the communication environment over time. We have confirmed the effect as a result of experiments in the room, so we explained the details. And we reported on the successful operation of the communication system and its effectiveness against communication interference.

## ACKNOWLEDGMENTS

The research is supported by Nonoichi City.

## REFERENCES

- [1] Hikaru Yabe, Shuto Ishikawa, Shunsuke Tomioka, Sho Tsukahara, Ryotaro Sakauchi, Mikiko Sode Tanaka, "Bus Location System with LoRa to Cover the Entire Nonoichi City," 2021 IEEE 3rd Global Conference on Life Sciences and Technologies (LifeTech), March 9-11, 2021.
- [2] T. Boshita, H. Suzuki and Y. Matsumoto, "IoT-based Bus Location System Using LoRaWAN," 2018 21st International Conference on Intelligent Transportation Systems (ITSC), 2018, pp. 933-938, doi: 10.1109/ITSC.2018.8569920.
- [3] P. Guan, Z. Zhang, L. Wei and Y. Zhao, "A Real-Time Bus Positioning System Based on LoRa Technology," 2018 2nd International Conference on Smart Grid and Smart Cities (ICSGSC), 2018, pp. 45-48, doi: 10.1109/ICSGSC.2018.8541282.
- [4] Shinichiro Kakuda, Yudai Yamazaki, Keita katagiri, Takeo fujii, Osamu Takyu, Mai ohta, and Koichi adachi, "Channel Allocation for LoRaWAN Considering Intra-System and Inter-System Interferences," IEICE-SR2021-12, pp.79-85, 2021-05-13.
- [5] Y. Ye and B. Wang, "RMapCS: Radio map construction from crowdsourced samples for indoor localization," IEEE Access, vol. 6, pp. 24224–24238, April 2018.
- [6] S. Bi, J. Lyu, Z. Ding, and R. Zhang, "Engineering radio maps for wireless resource management," IEEE Wireless Commun., vol. 26, no. 2, pp. 133–141, April 2019.
- [7] J. Haxhibeqiri, I. Moerman, and J. Hoebeke, "Low overhead scheduling of LoRa transmissions for improved scalability," IEEE Internet of Things J., vol. 6, no. 2, pp. 3097–3109, April 2019.
- [8] Z. Qin and J.A. McCann, "Resource efficiency in low-power wide-area networks for IoT applications," in Proc. IEEE Global Commun. Conf. (GLOBECOM), Singapore, pp. 1–7, Dec. 2017.
- [9] 920MHz-BAND TELEMETER TELECONTROL AND DATA TRANSMISSION RADIO EQUIPMENT, ARIB STD-T108 Version 1.3

# Localization Method for an Autonomous Cart as a Guard Robot

Yuya Sawano\* Yuto Nagai\*, Takayuki Suzuki\*\*, Ryoza Kiyohara\*\*†

\*Graduate School of Kanagawa Institute of Technology, Japan

\*\*Kanagawa Institute of Technology, Japan

**Abstract** - Some robotic cars have been develop. These small robotic cars are suitable for use as transport, guide, and patrol robots on campuses or in buildings. These small robots move autonomously using a location information system, obstacle detection system, and maps. We focus on a patrol system for finding suspicious but harmless people such as aged wanderer. We assume that the patrol robots have GPS for their location information system, laser imaging detection and ranging (LiDAR) for obstacle detection, an infrared sensor for human detection, and a Bluetooth device for authorized human detection. In this paper, we describe localization method for these small robotic cars.

**Keywords:** patrol, automated robotic car, LiDAR, Infrared sensor, detection of innocent suspicious person

## 1 INTRODUCTION

Numerous studies have recently been conducted on automated vehicles with many types of sensors, digital maps, and route guidance functions. Automated vehicles must avoid many types of accidents. Therefore, unmanned operated vehicles can operate in limited areas, and many studies on robotic vehicles that can be used within limited routes have been conducted [1–3]. However, new technologies for automated vehicles can reduce the many constraints placed on robotic vehicles. Therefore, some systems for robotic vehicles such as described in [1] have been developed. Such robotic vehicles are available for users within a limited area such as the inside of a campus or building.

The applications of such robot are as follows:

- (1) delivering services.
- (2) guidance services.
- (3) security services.

We are planning to introduce robotic vehicles on the campus of Kanagawa Institute of Technology to reduce costs. Through such an introduction, a delivery service can reduce its number of delivery and receiving personnel. Guidance services can use such vehicles to benefit their guests. Finally, security services can reduce their patrol costs.

In this study, we focus on security services such as patrols. The patrol service has the following objectives:

- (1) Detecting suspicious people who try to avoid a patrol (i.e., a deterrent).
- (2) Detecting and leading suspicious people who are not malicious.
- (3) Detecting suspicious objects.

However, it is difficult to send the exactly location information. Generally, GNSS is applied for these purposes for outdoor environments. However, there are multi-path problems. Therefore, in this paper, we propose the localization method for this system.

## 2 PATROL

### 2.1 Patrol in KAIT Campus

In this section, we describe the aims of the patrol used on our campus. Figure 1 shows the KAIT campus. There are many buildings on campus. Buildings no. 2 and no. 5 are only used for lectures. Other buildings have lecture rooms, personal rooms for professors, and rooms for office workers.

During the day, the main purpose of a patrol is finding suspicious objects included illegal parking. Because there are many students, teachers, office workers, and neighborhood residents on campus during the daytime, they are not required to go through an admission and confirmation process to enter the campus. However, all rooms except lecture halls are locked, and only students and teachers with approval can enter them.

During the nighttime, only teachers and students with permission can stay on campus. Therefore, a patrol service is required. The purposes of such a patrol were mentioned in section 1. Currently, security patrols are provided outside the campus buildings several times at night. Moreover, they patrol inside the buildings for other purposes. Therefore, they cannot conduct a patrol if a problem occurs.

Thus, we studied a way to introduce patrol robots on our campus. The robot works as a delivery and guide robot during the daytime. At night, the robot works as a patrol robot. The main purpose of introducing such robots are to reduce costs. Another purpose is a safe patrol. When finding a person who has fallen, the patrol robot might have to put on a gas mask. If the robots have a gas sensor, they can send such information to the control center. Figure 2 shows the goal of our study.

The control center controls the parking, charging, and communicating space through WiFi. There are many buildings and many objects between buildings (e.g., benches and trash bins). We are planning to put 5G equipment between each building. In these areas, the possibility of finding a human is greater than in other areas. Therefore, robots should be able to be controlled by the center using 5G.

It is difficult to find suspicious people. However, robots may encounter non-malicious people. Because, they do not think he is not illegal. Then, we hope that the robot can deter and find non-malicious people. Examples of non-malicious people include aged wanderers and children.

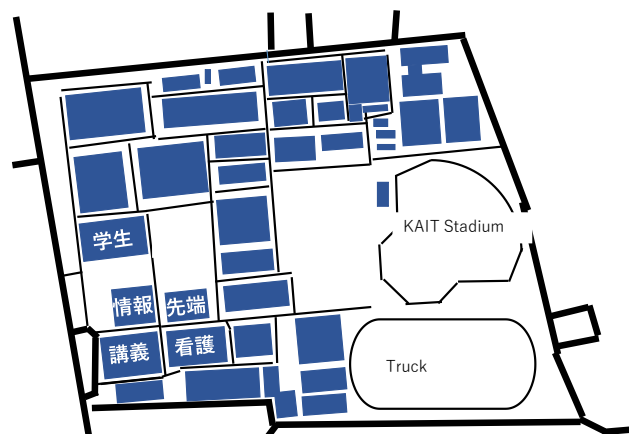


Figure 1: Map of KAIT campus

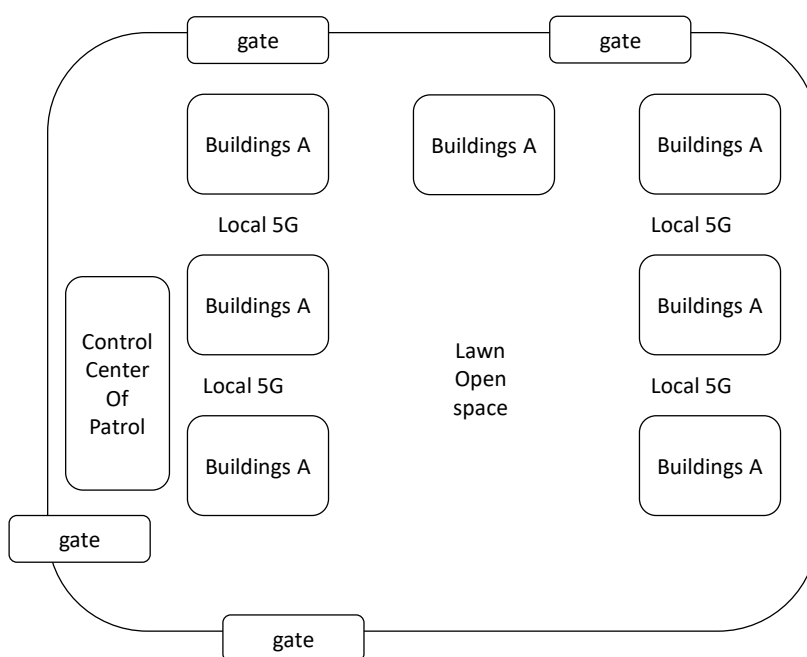


Figure 2: Image of goal of our study

## 2.2 Current Patrol

Currently, many patrol systems consist of human patrollers and monitoring cameras. These cameras show the surrounding area and information for security guards. Moreover, the camera is used for deterrence. However,

security guards must decide what action they should take for every case when applying this system. This means that the guards are constrained by time. Moreover, it is not easy to analyze the still images from monitoring cameras at night or under poor weather conditions.



### 3 ROBOTS

We are planning to introduce robots for many different purposes. The robots are based on small automated vehicles, and have many sensors:

- GPS
- 2D and 3D LiDAR for object detection
- Infrared camera

Moreover, robots can communicate through a network such as LTE, WiFi, Bluetooth, and 5G.

We assume the robots are a type of autonomous vehicle. Therefore, they have a campus map. Moreover, they have 3D LiDAR (VLP-16) [4] and GPS. In addition, they can update the map dynamically using their own 3D LiDAR and can compare the previous information with the current information.

Moreover, we assume there are no steps outside of the buildings on our campus. The robot speed is almost the velocity of a walking human.

### 4 RELATED STUDIES

There are many researches for localization. Main purpose of these research is to improve accuracy. [5] showed the improved method of under multi-path environment without the reference point such as RTK-RNSS. This method use the Infrared Omnidirectional Vision (IR-ODV).

[6] showed the method which use the LiDAR and Map. However, these methods require the additional devices..

### 5 BASIC EXAMINATION

We got many real data in our campus which data are by GNSS, BLE and LiDAR.

#### 5.1 GNSS

We measured the GNSS using by application [7] on the Android devices. At some places, there are many unacceptable error which are maximum 30 m. However, if the devices are far from building, there are acceptable error. Therefore, GNSS can be user the places which are far from buildings.

#### 5.2 BLE

There are many research about BLE with indoor environment. We assume the BLE points at the door. Therefore, if the device sense the BLE, we can decide the device is near the building or not. we can calculate by equation (1)(2).

$$RSSI = TxPower - 20 \times \log D \quad (1)$$

$$D = 10^{((TxPower - RSSI)/20)} \quad (2)$$

Then, we measured the BLE beacon at many places which are 5m or 10m from BLE beacon in our campus. Table 1 shows the setting information. Figure 3 shows the distances by RSSI data which are actual 5m points, Figure 4 shows the distances by RSSI data which are actual 10m points.

Table 1. BLE setting

standard	Bluetooth 4.0
Communication distance	50m
Interval	900ms
Txpower	-59dBm

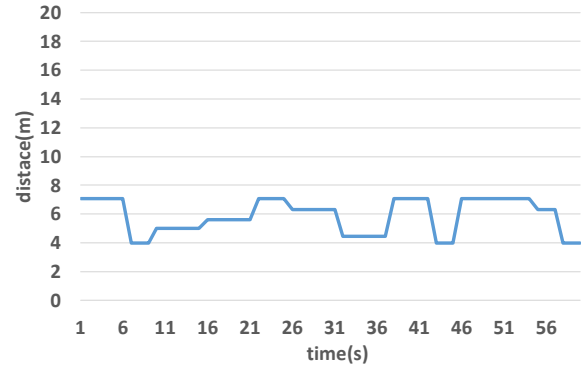


Figure 3. calculated distance at 5m point

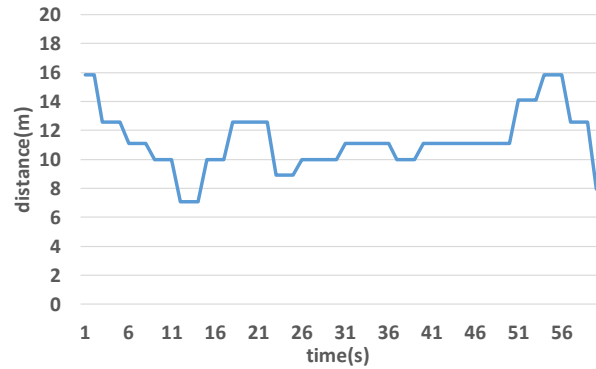


Figure 4. calculated distance at 10m point

From these results, if the distance is less than 5m, the information is used for localization. However, more than 5m, there are possibility of error.

#### 5.3 LiDAR

Our Robot has already RPLiDAR AiM8. Therefore, we measure this devices area of measurement. Figure 5 shows the result of around the devices. However, this experiment is executed at the indoor environment. This results shows that the Lidar can measure the distance from the buildings.

### 6 PROPOSED METHOD

we proposed the localization method which are combined GNSS, BLE and Lidar from the results of basic examination. If the robot is far from buildings, we use the GNSS. However, we cannot decide by only GNSS. Therefore, if the signals of BLE are low and the robot is far from buildings by Lidar, GNSS information should be used for localization.

If the signal of a BLE beacon is strong, the location of beacon is used for localization. If more than three signals of

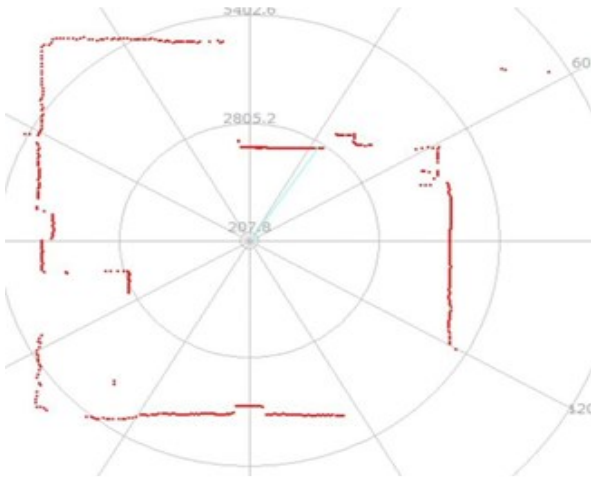


Figure 5. shape of our laboratory

beacons are detected, we calculate the location and the information is user for localization. Otherwise, map matching are applied the Lidar information.

## 7 EXPERIMENT

We will apply our proposed method at 26 points in our campus. Figure 5 shows the pointes

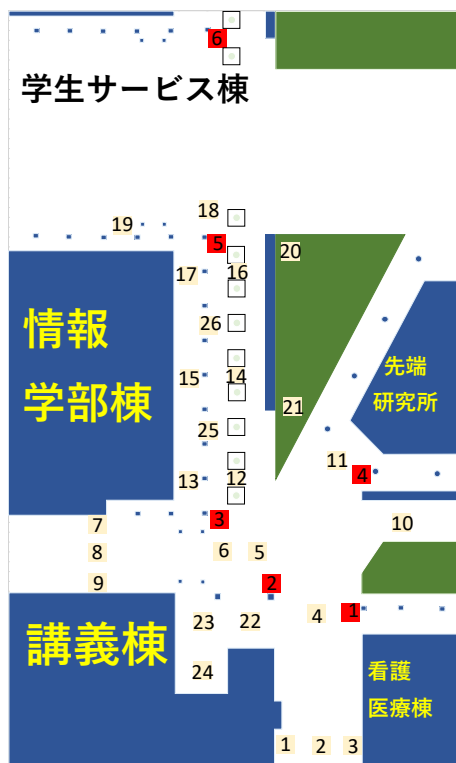


Figure 6 measured points

In the near future, we show the result of this experiments.

## 8 SUMMARY

We propose a localization method for the autonomous robots. That method is combined method of GNSS, BLE and LiDAR. We believe that the accuracy will be improved by our method.

## REFERENCES

- [1] SMP ROBOTICS: [https://smrobotics.com/security\\_robot/security-patrol-robot/](https://smrobotics.com/security_robot/security-patrol-robot/)
- [2] BIFUE USHIJIMA: <https://www.gov-online.go.jp/pdf/hlj/20181201/24-25.pdf>, Science and Technology, 2018
- [3] Security Robots on Patrol, <https://www.security-magazine.com/articles/89471-robots-on-patrol> Security Magazine
- [4] Velodyen: VLP-16, <https://velodynelidar.com/products/puck/>
- [5] Taro SUZUKI, Mitsunori KITAMURA, Yoshiharu AMANO, Takumi HASHIZUME ,\*
- “GNSS Precise Point Positioning Based on Multipath Signal,” Transactions of the Society of Instrument and Control Engineers, Vol.47, No.7, pp.399-405, 2012
- [6] Kentaro KIUCHI, Yoji KURODA,” Reduction of the accumulated error by the matching using LIDAR and Digital map, “No14-2 Proceedings of the 2014 JSME Conference on Robotics and Mechatronics.
- [7] GPSTest , ver3.8.4 , [https://play.google.com/store/apps/details?id=com.android.gpstest&hl=en\\_US&gl=US](https://play.google.com/store/apps/details?id=com.android.gpstest&hl=en_US&gl=US)

# A CyReal Approach to Sensor System Development

Kei Hiroi<sup>†</sup>, Akihito Kohiga<sup>‡</sup>, and Yoichi Shinoda<sup>‡</sup>

<sup>†</sup>Disaster Prevention Research Institute, Kyoto University, Japan

<sup>‡</sup>Japan Advanced Institute of Science and Technology, Japan  
hiro@dimsis.dpri.kyoto-u.ac.jp

**Abstract** - IoT devices are expected to enable inexpensive and easy measurement and collection of a wide range of environmental information, especially in the field of disaster prevention. Whereas, preliminary verification is difficult, because of a functional design that assumes their distributed deployment, the cost of development itself. Therefore, we develop a sensor system emulator with CyReal concept, which integrates the virtual device with the actual device, and federated with other simulators. This paper presents a prototype of our sensor system, and discusses a design for flexible integration, in order to support the development, update, debugging, and operation of sensor systems as a distribution network for disaster prevention information.

**Keywords:** sensor network, emulation system, sensor testbed

## 1 Introduction

IoT (Internet of Things) devices equipped with sensors and related technologies are expected to enable inexpensive and easy measurement and collection of a wide range of environmental information. In the field of disaster prevention, they are being utilized for various fields such as river observation and slope measurement, and are expected to be useful for collecting data at points where it has been difficult to figure out the condition of the environment.

Conventional environmental monitoring for disaster response is usually installed at a few vulnerable sites that require monitoring. The environmental observation had been carried out through a robust monitoring system using a leased network lines. However, due to the frequent and large-scale river floods in recent years, we face on pressing importance requiring a larger-scale monitoring network, even for small rivers and waterways that have not been monitored well in the past. An observation network using IoT has the potential to minimize the cost of implementation and operation. This implies that a large-scale observation system can be applied.

However, installation of a large number of sensor devices creates a new problem. IoT devices are expected to be used in urban areas and mountainous regions. Devices can be installed in large numbers to provide valuable measurements that have not been possible with conventional observation systems. Whereas, installing a large number of devices over a wide area makes it complicated to handle them and difficult to improve them by relocating them after installation. Especially in data measurement where the relationship between devices is meaningful, it is important to design functions and consider how to utilize them with distributed deployment. Preliminary

verification in this paper indicates to test the operation of the entire sensor network before installation, to find bottlenecks, and to consider solutions based on this functional design. By conducting preliminary verification before installing sensor devices, we can clarify data from the unique characteristics of sensor devices and the relationship between data from multiple sensor devices, as well as the transmission characteristics of data due to terrain and communication infrastructure, and use these results to design measurements that satisfy the purpose of use. Nevertheless, since a large number of devices are required, it is difficult to verify the functions in advance by preparing the necessary number of actual devices, which increases the cost of development itself and improving such as relocation. Although various sensor emulators have been developed to enable such preliminary verification, they are limited to verification of some functions such as network performance and device performance after using virtual devices.

In this study, we develop a sensor emulator system to support research, development, and operation of disaster prevention information collection and operation, and to enable preliminary verification assuming a specific installation environment and operational configuration. A sensor emulator is a virtualization technology that emulates the computers performance involved in sensor devices and their functions. In this research, we develop an emulator system that can exchange the virtual device with the actual device and can be federated with other simulators in order to support the development, update, debugging, and operation of sensor systems as a distribution network for disaster prevention information. By separating the Node and Sensor, and incorporating a connection mechanism between the virtual and physical devices for each of them, we can verify the functions of a large-scale sensor device. By enabling connection and verification not only with the virtual device but also with the physical device, it is possible to assume functions and communication environments that cannot be demonstrated with the virtual device, and it also facilitates connection to the cloud, thereby reducing operating costs and enabling multiple use of resources based on the assumption of an actual operating environment. Federation with other simulators can be possible through data exchange using the IoT linkage infrastructure. Through collaboration, data that assumes actual operation can be incorporated into the sensor, and preliminary verification, including operational forms such as disaster response based on data collection, becomes possible.

## 2 Related Works

Sensor emulator is a powerful tool that helps researchers and users to consider the design of sensor networks. An enormous amount of effort has been invested in developing emulators for various technologies related to sensor networks; communication protocols, computer processes, application software. As the number of sensors increases with the spread of IoT, the need to handle a large number of sensors has led to the development of a number of sensor virtualization technologies. SenaaS [1] is IoT virtualization framework to support connected objects sensor event processing and reasoning. This framework provide an ontology design by a semantic overlay of underlying IoT cloud and a policy-based service access mechanism in terms of semantic rules. Bose et al. [2] have presented resource abstraction at the sensor level on Sensor-Cloud infrastructure with virtualization of sensors for developing applications in various fields. This is a design for virtual sensor on cloud station. SenseWrap [3] is a middleware architecture providing virtual sensors as representatives for any type of physical sensors. This middleware supports sensor-hosted services and a standardized communication interface that applications can use without having to deal with sensor-specific details.

Wireless networks, an indispensable technology for sensor networks, are also subject to emulation. TOSSIM [4] focuses on simulating a wide range of network interactions. TOSSIM, which features a high fidelity and scalability, can capture network behavior while scaling to thousands of nodes. COOJA [5] is a sensor network emulator aimed at cross-level simulation, allowing simultaneous simulation at many levels of the system; sensor node platforms, operating system software, radio transceivers, and radio transmission models. Many other emulators for sensor network verification have been researched and developed, such as EmStar [6], Avrora [7], and J-Sim [8]. These sensor emulators have been developed on the premise of wireless sensor networks. Recently, based on recent developments in low power wide area networks, including the rise of Long Range (LoRa) technology, a LoRa Coverage Emulator [9] has also been developed. This LoRa emulator consists of a transmitter and receiver and provides a reliable network coverage estimation based on the LoRa network design framework.

There are many researches on emulators from Hardware / Software point of view as well as network. The Freemote Emulator [10] is an emulator for developing software for nodes. It provides developers with a system architecture in several layers: Physical, Data Link (MAC), Routing and Application. Similarly, ATEMU [11] is a well-known sensor network emulator with a lot of contributions. A unique feature of ATEMU, which can operate on different application/hardware platforms, is its ability to simulate a heterogeneous sensor network. ATEMU emulates the processor, radio interface, timers, LEDs and other devices. Kasprwicz et al. [12] focused on CCD sensors as devices and developed a hardware emulator to speed up and streamline post-assembling tests and debugging. Furthermore, research on emulators has also focused on commoditization, with an emphasis on lightweight and small IoT systems. Brady et al. [13] developed an em-

ulator for an IoT environment using the popular QEMU system emulator to build a testbed of inter-connected, emulated Raspberry Pi devices. The effort to emulate functionality extends to applications that anticipate not only specific devices, but also power supply and utilization. Deda et al. [16] have designed a battery emulator/tester system to reduce development time for developing and testing of high-voltage power supply systems. Abrishambaf et al. [14] have developed a laboratory emulation model for energy scheduling in an agriculture system using real nodes.

The conventional emulators so far can be said to be emulators that specialize in a certain function of the computer. By using these technologies, hardware, software, and network can be emulated in an integrated manner. However, with the evolution of IoT, we have to treat power supply, heterogeneous sensor devices, and network devices. We need to provide an operator-friendly verification environment. The benefits of the IoT have increased the opportunities for sensing technologies to be more readily available to a wider user. And this advantage means that sensor networks can be built more widely and in more places than ever before. Verifying the required functionality with a few emulators may provide the required results. Although, for actual operators whose work is on the use of application services, it is very difficult to verify the functions in isolation. Nonetheless, emulators that require special technology are considered to have little affinity with this kind of operators. Also, for operators, when considering many things that could not be verified with previous emulators, such as geospatial and distributed installation for business efficiency, these requirements are not considered with previous emulators.

Therefore, we defined the requirements for a sensor emulator system to meet these requirements as follows.

- The sensor system emulator should be compatible with a real system.
- The sensor system emulator should be able to handle a large number of sensors.

To meet the above requirements, we develop an emulator that incorporates the concept of CyReal.

## 3 Sensor Systems by CyReal Emulator

### 3.1 Overview

The sensor emulator system developed in this paper enables preliminary verification of IoT devices based on the assumption of their specific installation environment and operational configuration. Based on the concept of CyReal, this sensor emulator system is designed to support the development, update, debugging, and operation of sensor systems, so that the virtual machine and the actual machine can be exchanged and can be federated with other simulators.

### 3.2 CyReal Approach

CyReal is intended to be an entity that plays an intermediate role in the concept of digital twin [15]. CyReal enables

flexible replacement and integration of real and virtual entities, such as computer systems, people, and environments. The configuration based on CyReal strongly promotes the digital twin of disaster prevention IT systems. Digital twin refers to the creation of digital objects to be handled in the real world and computer systems. This is the concept of debugging various properties on the created twin, in the case of computers, and is an important concept in the AI world.

Namely, the configuration on the left in Figure 1 is one in which all subsystems are configured by simulators, and the subsystems are connected via a federation platform. On the other hand, the various simulations and systems that configure this platform can all be replaced with real things, real systems. The configuration on the right side of Figure 1 shows a situation in which the subsystems are all real and in actual operation, such as people, natural phenomena, and a real disaster prevention information system. These two are in a digital twin relationship. Then, we have further extended the digital twin concept to allow real systems and simulators to be integrated in a subsystem (The center of Figure 1). This concept is beginning to be called CyReal, as an integration of Cyber and Real.

### 3.3 CyReal Sensor System

The sensor system we are developing is configured in accordance with the CyReal concept. The sensors are connected as subsystems in Figure 1. The sensors can be replaced by real or virtual ones. We expect that this CyRealization allows the subsystems to work in various ways and the system to be used for various purposes. That is, depending on the system we replace, this configuration can be transformed into a system with various purposes.

If the system is entirely composed of simulators, it will work as a disaster management IT simulator. For example, new analysis technologies and simulators can be connected to the disaster prevention IT simulator, and all simulators can be run based on data from past disasters. Since data is difficult to collect in disaster research, evaluating the performance of analysis technologies and simulators is a difficult and costly task. By creating this disaster prevention IT simulator, we have the prospect of providing an evaluation environment and facilitating development. Alternatively, if the subsystems are replaced with real ones (i.e., if the sensor system is replaced with a real one), it becomes a test bed that can be used for research and development and performance evaluation of the sensor system.

This configuration eliminates the limitation of devices, simulators, and systems that can be verified, which is a requirement of this paper. In other words, we can connect not only sensor devices, but also geospatial and operational simulations to verify the functionality and effectiveness of the system in actual operations. In order to develop a system that can flexibly switch between these three modes, we have embarked on a sensor system as one of the proofs of concept for the digital twin and CyReal that bridges the gap between Real and Virtual.

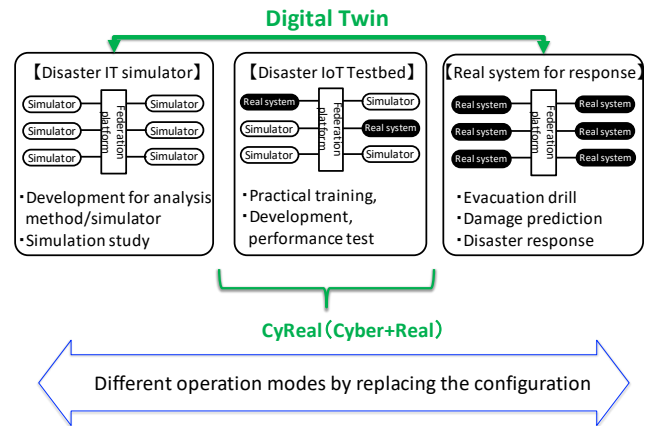


Figure 1: CyReal Approach for Sensor System

### 3.4 Significance of CyReal Sensor System

This sensor system is not a simple sensor virtualization. We have designed the system to anticipate the actual data utilization of the sensors. Conventional sensor virtualization has only simulated the data and functions involved in sensing. On the other hand, we not only simulate the data, but also develop the sensor device and its environment as an instance.

Our expected applications are as follows. Data missing is an unavoidable problem in large-scale sensor deployments. It is difficult for a system that only simulates data to represent the handling of such missing data. Our system design is able to produce data by incorporating not only the sensor device itself, but also the external environment, such as the wireless environment and the conditions of the location where the sensor is installed. This makes it possible to represent fluctuations in the data due to the influence of the external environment (Figure 2). In addition, data diversity can be expressed by integrating with external simulators, such as simulators for operations and simulators for natural phenomena. This is the advantage of an emulator that can replace and integrate actual and virtual nodes and sensors. Such a sensor system has many uses in verifying operations, but it also has many challenges. This paper shows proof-of-concept for a sensor system according to this concept, investigates its performance when running virtual sensors on a large scale, and consider the challenges of implementation.

### 3.5 Structure of CyReal Sensor System

Based on the CyReal concept, we have developed a sensor emulator system that can connect and verify not only virtual devices but also real devices. A sensor emulator is a kind of virtualization technology that enables preliminary verification of sensor-related systems, such as network performance, device performance, and computer processing performance. The sensor emulator is used for functional verification using virtual sensor devices. Conventional technologies and prior research to date have only supported virtual sensor devices, so that it is difficult to conduct verification based on actual operational environments. Recent IoT devices and related technologies are said to make it possible to distribute sensors over a wide area and to measure and collect data easily

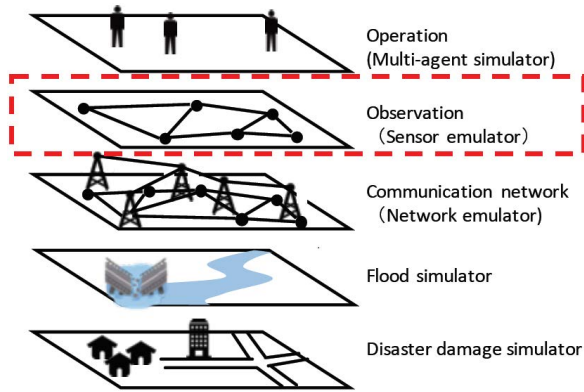


Figure 2: Federation with Simulators for External Environment

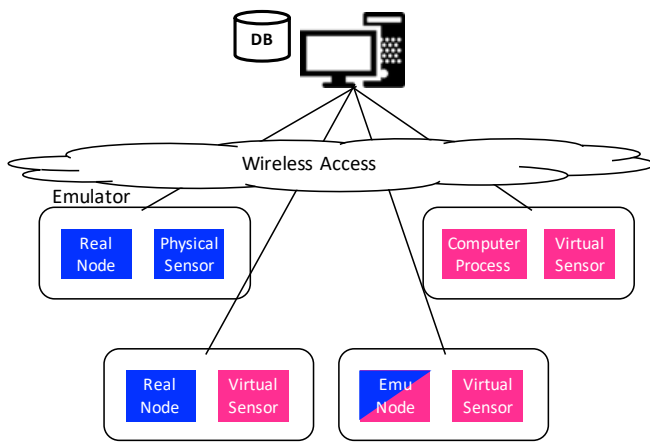


Figure 3: Structure of CyReal Sensor System

and inexpensively. However, in reality, it is difficult to design functions assuming distributed deployment and to prepare the necessary number of actual devices for functional verification in advance, and this has the disadvantage of leaving the decision to the user. Therefore, we have attempted to develop a CyReal sensor system to support research and development and operations related to the collection and distribution of disaster prevention information. In Figure 3, Sensor refers to a measurement device and Node refers to a computer that processes and communicates measurement data. Here, both sensor and node are designed to be interchangeable between actual and virtual machines. Furthermore, the system can be communicated with other simulators. This is based on data from past disasters, and is intended to be extended to verification based on actual situations such as power outages, device failures, and network disconnection.

## 4 Discussion and Conclusion

It is very difficult to prepare properly preliminary verification for a large-scale IoT environment, in consequence, we have not benefits from IoT technologies. To address this problem, this paper proposes a sensor system emulation environment based on the CyReal concept that integrates real and virtual systems. We aim to support research, development, and operation of disaster prevention information collection

and operation, and to enable preliminary verification assuming a specific installation environment and operational configuration. We expect to further develop the sensor system emulator in this paper and for verification by federating with other simulators in order to improve the efficiency of disaster response.

This paper developed a prototype of a sensor system that integrates real and virtual systems. The result of the experiments on the prototype showed the number of devices that can be used as a guide for larger scale in a general use environment, which is sufficient to operate the number of sensors currently used for disaster response. We plan to define CyReal-ness and improve the system to allow selective use of Real / Virtual at several layers: hardware, algorithms, and data. We expect preliminary verification for IoT systems from both functional and utilization aspects, including more realistic verification such as data missing.

Our objective is to develop a sensor emulator that enables preliminary verification of the behavior of the entire sensor network before installation to identify bottlenecks and determine how to resolve them, based on the premise of functional design. It is intended to be used to design measurements that satisfy the purpose of use by clarifying data due to the unique characteristics of sensor devices, relationships between data from multiple sensor devices, and data transmission characteristics due to terrain and communication infrastructure. In this paper, we have shown that Physical/Virtual Sensor/Nodes can be integrated in a CyReal Sensor System that work on a typical computational environment. In order to use the CyReal Sensor System for functional design in the future, it is necessary to add functions for various conditions such as network conditions, power consumption, and environmental exposure due to the installation location, and then integrate these functions with the physical and virtual conditions. Even if we consider only the network, the variety of infrastructures available makes it necessary to consider a complex configuration to build these functions on the sensor device. Therefore, we are trying to solve this problem by developing a simulation layer that enables CyReal of network and power on a separate layer from the sensor device, and federating these simulations. Moreover, we plan to develop a verification environment for distributed installation from the perspective of geographical characteristics of sensor installation and utilization through MQTT-based federation with other simulators for external environment.

## Acknowledgement

This work was supported by JST, PRESTO Grant Number JPMJPR2036, Japan.

## REFERENCES

- [1] A. Sarfraz, M. MR Chowdhury, J. Noll, Senaas: An Event-driven Sensor Virtualization Approach for Internet of Things Cloud, In 2010 IEEE International Conference on Networked Embedded Systems for Enterprise Applications, pp. 1–6 (2010).



- [2] S. Bose, A. Gupta, S. Adhikary, N. Mukherjee, Towards a Sensor-cloud Infrastructure with Sensor Virtualization, In the Second Workshop on Mobile Sensing, Computing and Communication, pp. 25–30 (2015).
- [3] P. Evensen, M. Hein, SenseWrap: A Service Oriented Middleware with Sensor Virtualization and Self-configuration, In 2009 IEEE International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), pp. 261–266 (2009).
- [4] P. Levis, N. Lee, M. Welsh, D. Culler, Tossim: accurate and Scalable Simulation of Entire Tinyos Applications, In Computer Communications and Networks, International Conference on Embedded networked sensor systems, pp. 126–137 (2003).
- [5] F. Osterlind, A. Dunkels, J. Eriksson, N. Finne, T. Voigt, Cross-level Sensor Network Simulation with Cooja. In 31st IEEE conference on local computer networks, pp. 641–648 (2006).
- [6] L. Girod, N. Ramanathan, J. Elson, T. Stathopoulos, M. Lukac, D. Estrin, Emstar: A Software Environment for Developing and Deploying Heterogeneous Sensor-actuator Networks, In ACM Transactions on Sensor Networks (TOSN), 3(3), pp. 1–34 (2007).
- [7] B. Titzer, D. Lee, J. Palsberg, Aurora: Scalable Sensor Network Simulation with Precise Timing, In IEEE Fourth International Conference on Information Processing in Sensor Networks (IPSN '05), pp. 477–482 (2005).
- [8] A. Sobeih, J.C. Hou, L. Kung, N. Li, H. Zhang, W. Chen, H. Tyan, H. Lim, J-Sim: A Simulation and Emulation Environment for Wireless Sensor Networks, IEEE Wireless Communications, 13(4), pp. 104–119 (2006).
- [9] B. Al Homssi, k. Dakic, S. Maselli, H. Wolf, S. Kandeeppan, A. Al-Hourani, IoT Network Design Using Open-Source LoRa Coverage Emulator, IEEE Access, No.9, pp. 53636–53646 (2021).
- [10] T. Maret, R. Kummer, P. Kropf, J. F. Wagen, Freemote Emulator: A Lightweight and Visual Java Emulator for WSN, In International Conference on Wired/Wireless Internet Communications, pp. 92–103 (2008).
- [11] J. Polley, D. Blazakis, J. McGee, D. Rusk, J. Baras, Atemu: a Fine-grained Sensor Network Simulator, In Sensor and Ad Hoc Communications and Networks, pp. 145–152 (2004).
- [12] G. Kasproicz, L. Mankiewicz, K. T. Pozniak, R. S. Romaniuk, S. Stankiewicz, G. Wrochna, Hardware Emulator of the High-resolution CCD Sensor for the Pi of the Sky Experiment, In Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments 2007, Vol.6937, pp. 693–708 (2007).
- [13] S. Brady, A. Hava, P. Perry, J. Murphy, D. Magoni, A. O. Portillo-Dominguez, Towards an Emulated IoT Test Environment for Anomaly Detection using NEMU, In 2017 Global Internet of Things Summit (GIIoTS), pp. 1–6 (2017).
- [14] O. Abrishambaf, P. Faria, Z. Vale, Laboratory Emulation of Energy Scheduling in an Agriculture System. In 2020 IEEE/PES Transmission and Distribution Conference and Exposition (T&D), pp. 1–5 (2020).
- [15] S. Boschert, R. Roland, Digital Twin—The Simulation Aspect, Mechatronic futures, Springer, pp. 59–74 (2016).
- [16] S. Deda, A. Eder, V. Mhetre, A. Kuchling, R. Greul, O. Koenig, Designing a Battery Emulator/Tester from Scratch to Prototyping to Automated Testing within a HIL Digital Twin Environment, In International Exhibition and Conference for Power Electronics, Intelligent Motion, Renewable Energy and Energy Management, pp. 1–8 (2020).
- [17] Z. Ye, F. Hu, L. Zhang, Z. Chu, Z. O'Neill, A Low-Cost Experimental Testbed for Energy-Saving HVAC Control Based on Human Behavior Monitoring, International Journal of Cyber-Physical Systems (IJCPS), 2(1), pp. 33–55 (2020).
- [18] P. Evensen, M. Hein, Sensor Virtualization with Self-configuration and Flexible Interactions, In the 3rd ACM International Workshop on Context-Awareness for Self-Managing Systems, pp. 31–38 (2009).
- [19] H. Debnath, N. Gehani, X. Ding, R. Curtmola, C. Borcea, Sentio: Distributed Sensor Virtualization for Mobile Apps, In 2018 IEEE International Conference on Pervasive Computing and Communications (Per-Com), pp. 1–9 (2018).
- [20] Z. Wang, M. Liu, S. Zhang, M. Qiu, Sensor Virtualization for Underwater Event Detection, Journal of Systems Architecture 60.8, pp. 619–629 (2014).





# Proposal of an efficient downward communication method for a large-scale data collection system using MQTT

Fuya Aoki, Koichi Ishibashi, and Tetsuya Yokotani

Kanazawa Institute of Technology, Japan

c6101709@planet.kanazawa-it.ac.jp, {k\_ishibashi, yokotani}@neptune.kanazawa-it.ac.jp

**Abstract** - In recent years, MQTT has been attracting attention as an important means of communication in the deployment of the IoT. MQTT is being researched and developed from various angles to realize IoT services employing MQTT. However, there are some challenges such as scalability, interoperability and so on. In this paper, we propose a method for efficient downward communication for a large-scale data collection system. The proposed method consists of connection management by a server and downward transfer by tunneling. That is, the server manages information of a broker to which a sensor is connected, and transfers a message to the specific broker by utilizing tunneling technique when the server notifies data to the sensor. Therefore, the proposed method enables a reduction in the amount of memory for connection management in the specific broker, and avoids an increase in communication traffic due to memory overflow. We evaluated a communication traffic amount of the proposed method and experimentally confirmed the feasibility of data transfer by using the tunneling technique.

**Keywords:** IoT, MQTT, Large-scale data collection, Wireless sensor network, Connection management

## 1 INTRODUCTION

In recent years, IoT (Internet of Things) devices have been spreading rapidly, and not only personal computers, smartphones, and other Internet-connected terminals, but also home appliances, automobiles, factory equipment, street lights, and other “things” have been connected to the network, becoming indispensable not only in our daily lives but also in various industries. The number of IoT devices in the world is increasing and is expected to reach several hundred billion in the next few years [1]. In this context, it is desirable for IoT services to collect real data from a wide variety of IoT devices over a wide area to collect real data from a specific device, and to notify control data to a specific device when needed. To realize such an IoT system, research and development from various aspects such as the construction of a wireless sensor network (WSN) for accommodating tiny IoT devices and an IoT core network consisting of gateways for WSNs have been advanced [2][3][4][5]. In addition, ISO/IEC JTC1 /SC41 classifies use cases of IoT services, summarizes the requirements of a communication platform from the perspective of communication styles and QoS(Quality of Service), and proposes an IoT data exchange platform for various IoT

services to reduce communication traffic compared with an IoT system constructed on conventional networks.

In the field of IoT, a communication protocol called MQTT (Message Queuing Telemetry Transport) has attracted attention as an important communication method for data monitoring. MQTT is an asynchronous, lightweight communication protocol that consists of a publisher to send messages, a subscriber to receive messages, and a broker to mediate messages. In addition, the function of a data receiver attempting to receive data from the system by subscribing and the function of the sender notifying the data by publishing operate independently (Figure 1). However, its application to a large-scale system over a wide area is a challenging issue because the current MQTT specification defines the operation by a single broker.

In this paper, we propose a method to realize efficient downward communication in a large-scale data collection system consisting of a large number of IoT devices, such as a smart street lighting system. In Section 2, we summarize the issues in large-scale data collection systems; in Section 3, we show related research on building large-scale IoT systems using MQTT. In Section 4, we propose a method for efficient downward communication, and in Section 5, we show an evaluation of a traffic amount of the proposed method and experimental results to confirm the data transfer, based on a tunneling technique by utilizing the bridge function implemented on a MQTT software. Lastly, Section 6 provides our conclusion.

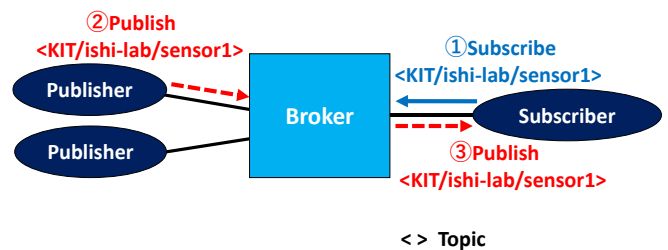


Figure 1: MQTT (Message Queuing Telemetry Transport).

## 2 CHALLENGES IN LARGE-SCALE DATA COLLECTION SYSTEMS

As an example of a large-scale data collection system, we focus on a smart street lighting system. Recent advancement in the smart street light with controllable LED and communication method has enabled the development of the remote-control system for managing and maintaining the street lighting devices on a central server [6][7]. The system

adopted for the street lighting has become an important research area because of the potential for efficient maintenance of lighting devices and provision of additional and novel services. For example, by visualization of state and keep-alive of a street lighting device on the central server, it is possible to reduce energy consumption based on optimal control of the dimming of light in consideration of the influence of buildings, trees, the weather conditions, and so on. In addition, it is also possible to improve the efficiency of device maintenance as well as reduction of maintenance costs due to personnel expenses.

Moreover, additional services are actively being considered, such as visual navigation. The visual navigation enables the smooth guidance of people by remotely controlling the lighting color and blinking frequency of street lights in the instance of a disaster or during a specific event, and contributes to conserving a safe and secure society. Another candidate of novel service, the visualization of sensor information by environmental monitoring, is discussed. This service collects environmental information from environmental sensors installed on a poll with street lights that are ubiquitous everywhere and provides environmental information such as temperature and humidity (Figure 2). In the smart street lighting system, small data such as sensor information are exchanged between a central server and sensors such as smart street lighting devices, and the overhead of resolving IP addresses and transferring data becomes large when using conventional IP address-based communication. Therefore, the application of MQTT, a communication protocol for the IoT with small overhead and lightweight, is effective. Here, a smart street lighting system is a system consisting of hundreds of thousands of sensors. Therefore, the scalability of the broker becomes an issue in a system that utilizes MQTT. In other words, when considering data collection by a server, data from many sensors will be concentrated on the server, and the brokers connected to the server will be overloaded.

It is possible to distribute the load of the broker by introducing multiple brokers, each of which can accommodate nearby sensors. However, this approach requires the registering of information, of all the sensors to the broker connected to the server to support the downward communication from the server to the sensors. Therefore, challenges of increasing the amount of memory on the brokers remain. Another solution is that the broker that accommodates each sensor connects to the server independently. However, the challenges of increasing the number of connections required by the server still remains.

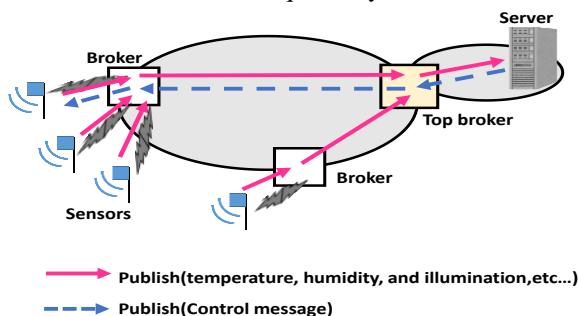


Figure 2: Smart street lighting system.

### 3 RELATED WORK

Regarding the realization of IoT systems using MQTT, some research has been studied on performance evaluation and proposed cooperation mechanisms with multiple brokers to deploy a large-scale system. For example, in [8][9], the performance evaluation of various implementations of MQTT against brokers has been studied. According to the benchmark in [8], the performance of a MQTT broker depends on the number of MQTT clients and a computer resource implemented a broker function. The large-scale data collection system for smart street lighting systems, which is the target of this study, remains a challenge to realize with a single broker.

Research on MQTT systems with multiple brokers has also been conducted in many places [10][11][12]. In [10], MQTT-ST is proposed to construct a distributed network of multiple MQTT brokers. MQTT-ST consists of a spanning tree of brokers in the network and shares messages among multiple brokers to cope with failures. However, when considering communication for data collection, it is expected to be effective, but when considering communication for data notification from the server, there are problems such as the broker being overloaded through the subscription of many sensors. In [11], a scalable and low-cost MQTT broker clustering system is proposed to handle a huge amount of IoT devices. In this clustering system, MQTT clients and multiple MQTT brokers are connected by a load balancer that distributes and balances the network traffic to the MQTT brokers. Therefore, compared to a single broker, the load on each broker is reduced, the throughput of the entire clustering system is increased, and the CPU utilization of each broker is reduced. In [12], MQTT brokers are placed at each network edge to handle data with the characteristic of “edge heavy,” where objects at the network edge of the IoT environment generate a large amount of data. To coordinate these multiple MQTT brokers, they are proposing a new mechanism called ILDM. An ILDM node placed between a broker and a client not only relays MQTT clients and brokers like a proxy but also can connect to other ILDM nodes, allowing multiple brokers to work together. As shown in [11][12], the deployment of a system using multiple brokers is being considered in many places to build a large-scale system. However, there is still the problem of increasing the traffic of the whole system.

Thus, research has been conducted on the realization of large-scale IoT systems using MQTT. However, according to the performance evaluation of a single broker, it is impossible to realize a large-scale data collection system with a single broker. In addition, through research into the collaboration of multiple brokers, it has been found that, when considering the communication for data notification from the server, the brokers become overloaded due to subscription from many sensors. So, the problem of increasing the traffic in the whole system, composed of multiple brokers, is still a challenge. Therefore, in this paper, we propose a method to construct a system, with multiple brokers, that facilitates efficient communication in the downward direction (i.e., the data notification from the server) without increasing the traffic.

## 4 PROPOSED METHOD

In this paper, we propose an efficient downward communication method for a large-scale data collection system using MQTT, targeting large-scale data collection systems such as smart street lighting systems. The target system consists of multiple brokers that IoT devices as sensors connect to at each point and a top broker that servers connect to, and it provides functions of data collection and data/command notification from sensors or a server by publish/subscribe using MQTT. The proposed method consists of connection management by a server and downward transfer by tunneling (Figure 3). That is, in the proposed method, the server manages the information of a broker, to which a sensor is connected, and transfers, by utilizing a tunneling technique, messages to the top broker that includes data to the sensor and information for the sensor managed by the server when the server notifies data to the sensor.

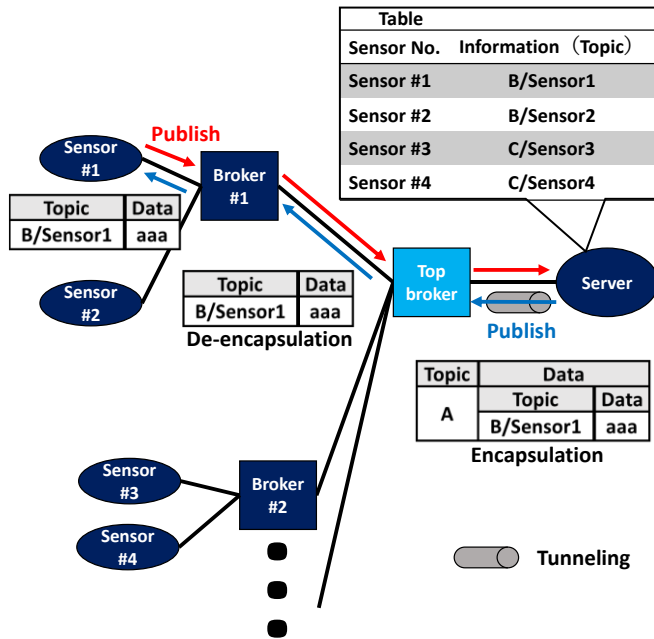


Figure 3: Connection management by the server and downward transfer by tunneling.

### 4.1 Connection management by server

First, we describe connection management by the server. Connection management, by the server, reduces an amount of memory on the top broker for the management of topics due to path concentration on the top broker for communication from the server to sensors. In this management method, the server manages the information of the sensors, which means the brokers to which the sensors are connected. For example, when the sensor #1 is connected to the broker #1, as shown in Figure 3, sensor #1 notifies the topic managed by the broker #1, that is, "B/Sensor1", as information of the sensor #1, to the server. Similarly, the sensor #3 notifies the topic managed by the

broker #2, "C/Sensor3", to the server. According to information notified from sensors, the server maintains a binding table for a relationship between a sensor and a broker to which the sensor connects.

### 4.2 Transfer by tunneling

Next, we explain a method for the transfer of notification data from the server to a sensor by tunneling. When the server notifies data to a sensor, the server transmits tunneling data, which includes data to the sensor and a topic for notification to the sensor. It is noted that the topic is equivalent to information for the sensor managed by the server.

For example, when the server notifies a piece of data, "aaa," to the sensor #1, the server publishes a message with topic "A" which includes the topic of the broker "B/Sensor1" in Figure 3, and the actual data to be sent, that is "aaa". The topic "B/Sensor1" means the destination of the actual data from the server's point of view. It is noted that the topic "A" is provided for the server to publish a tunneling message. When the top broker receives the message in the topic for the tunneling, the top broker extracts the data part of the received message and the original topic, and transfer the data to the destination broker according to the extracted topic.

### 4.3 Sequence of downward communication

Figure 4 shows a sequence for the downward communication of the proposed method. Since the server manages information of the sensor in the proposed method, at the initial stage, each sensor needs to register the information of the broker to which it is connected. That is, the sensor notifies information (which is the topic of the broker that the sensor connects to) to the server, and subscribes a topic "B/Sensor1" to the broker which it connects to, in order to receive data from the server at any time. Following this, when the server notifies a piece of data, "aaa," to the sensor, the server publishes a message with topic "A" which includes the topic of the broker "B/Sensor1" and the actual data "aaa". The topic "A" is a topic which the top broker provides to transfer received data to an encapsulate topic. When the top broker receives messages at topic "A", the top broker extracts the data part of the message and the encapsulated topic "B" from the received message, and transfers (publishes) the message to the corresponding destination broker according to the extracted topic. When the broker receives this message, it transfers the message to the subscribed sensors. Thus, the proposed method allows the top broker to transfer messages from the server to the desired broker without holding sensor information, thus reducing the amount of memory required on the top broker. In addition, it is said that the top broker functions as a proxy for a publisher.

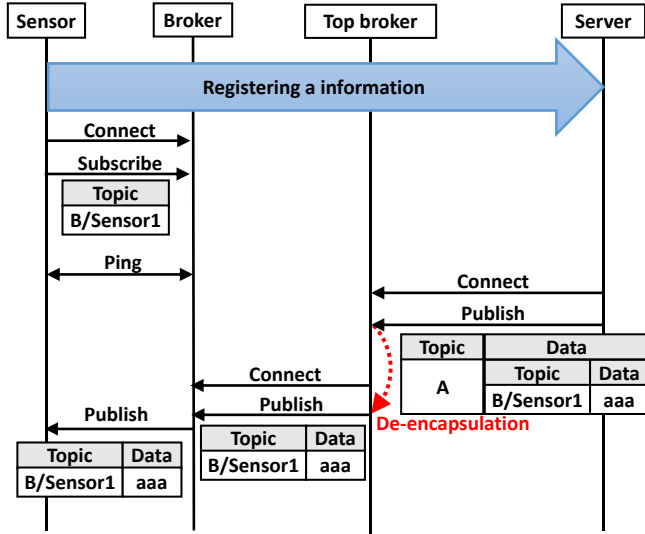


Figure 4: Communication sequence of the proposed method in the downward direction.

## 5 EVALUATION

To confirm the effectiveness of the proposed method in downward communication, we compared traffic generated when using the proposed method and when using a conventional method with a single broker on a large-scale data collection system using MQTT, based on simplified analysis models. In addition, to check the feasibility of the tunneling technique on the proposed method, we experimentally confirmed the data transfer from the server to a sensor by way of the top broker by utilizing the bridge function implemented on a MQTT software.

### 5.1 Evaluation of communication traffic

First, we compared the communication traffic in the IoT core network of the IoT system on both the proposed method and the single broker approach. On evaluation of the communication traffic of the IoT system by a single broker, it is assumed that a “publish” message with an unknown topic is broadcasted when a memory overflow occurs in the broker due to an increase in the topics subscribed by connected sensors. According to [8], it is assumed that memory overflow occurs when the number of subscribed topics from sensors reaches 100,000 or more. The configuration of the IoT system applied on the proposed system is shown in Figure 5 and the configuration of the IoT system by the single broker is shown in Figure 6. The sequence of the proposed method is shown in Figure 7, and the sequence of the IoT system by a single broker is shown in Figure 8. Figure 8 also shows that a “publish” message with an unknown topic for the corresponding topic, is broadcasted to each sensor in the event of a memory overflow in the broker. Equations (1) and (2) show traffic due to “publish” messages ( $T_1$ ) in the IoT core network on the IoT system by a single broker and traffic due to “publish” messages ( $T_2$ ) in the IoT core network on the IoT

system adopting the proposed method, based on simplified analysis models. That is, it is focused on the traffic due to “publish” messages, and it is assumed that the server fairly notifies a notification data to all sensors using MQTT. Where  $n$  is the number of sensors,  $p$  is the number of messages for publication from the server to a sensor, and  $p=1$ , in this simple analysis, in the first step. In the case of the IoT system using a single broker, when the number of connected sensors exceeds 100,000, the broker broadcasts a message to multiple sensors due to memory overflow.

$$T_1 = \begin{cases} n \times p & (n \leq 100000) \\ \{100000 + (n - 100000)^2\} \times p & (n > 100000) \end{cases} (1)$$

$$T_2 = n \times p (2)$$

Figure 9 shows a comparison of the amount of communication traffic through the number of messages. Although in the IoT system adopting a single broker, the amount of communication traffic increases rapidly when the number of sensors exceeds 100,000, due to the influence of memory overflow, the amount of communication traffic in the IoT system utilizing the proposed method increases linearly according to the number of sensors. Therefore, it is confirmed that the proposed method is effective when the number of sensors is large.

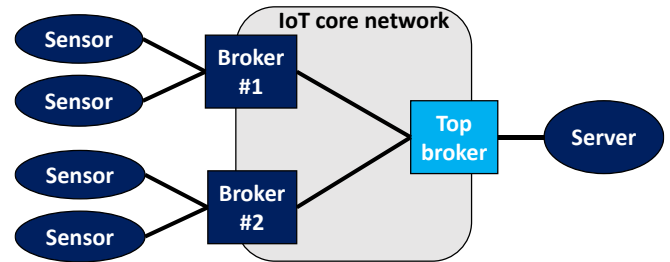


Figure 5: Configuration of the IoT system adopting the proposed method.

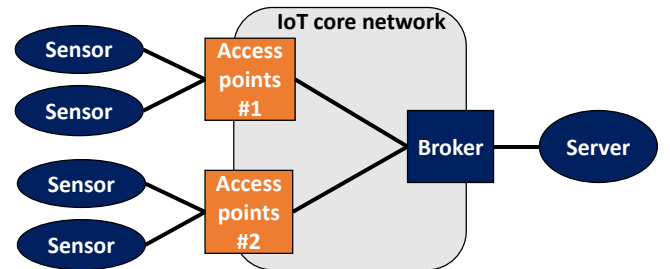


Figure 6: Configuration of an IoT system by a single broker.



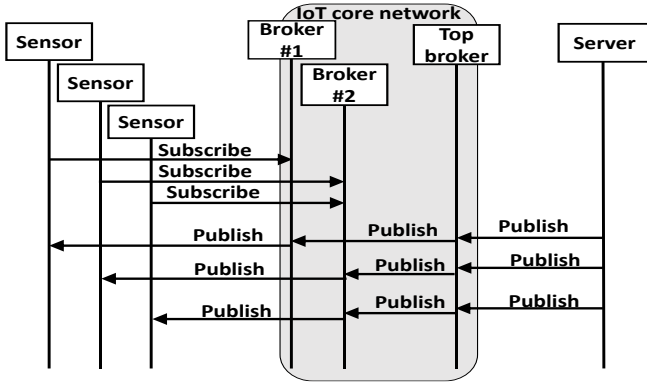


Figure 7: Sequence of evaluation of the proposed method.

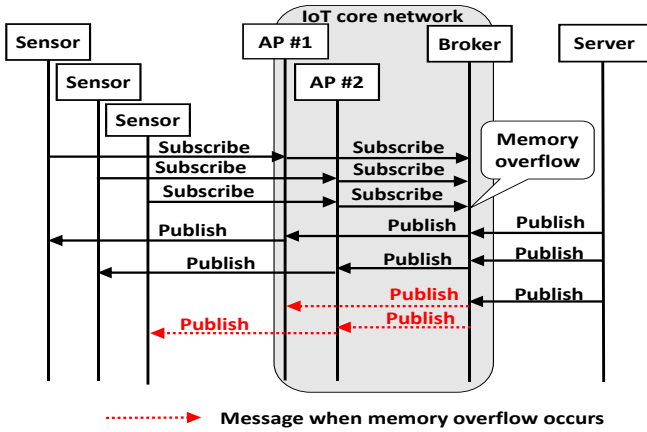


Figure 8: Single broker evaluation sequence.

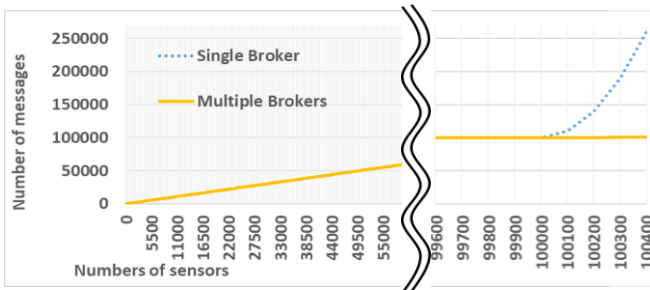


Figure 9: Comparison of the number of messages related to MQTT on the IoT core network.

## 5.2 Experimental verification of tunneling technique

To confirm the feasibility of the application of the tunneling technique within the proposed method, we constructed a system consisting of multiple brokers and verified the function of transferring from a publisher to a subscriber, using MQTT's bridge function, by way of a specific broker being equivalent to a top broker. It is noted that the bridge function of Mosquitto [13], which is MQTT software, is applied as a MQTT's bridge function. The bridge function is a function to share "publish" messages for a specific configured topic among multiple brokers. Figure 10 shows the experiment environment which consists of six Raspberry Pi units, with one top broker, two brokers, two sensors, and one server. To confirm that messages are

automatically transferred to each sensor via a broker, we monitored packet flows to the entrance on the server-side and exit on broker #1 and broker #2 sides of the top broker using Wireshark. The balloon in Figure 10 describes the configuration of the bridge of the top broker. Based on the topic of the message published from the server to the top broker, the top broker is configured to transfer the message to the broker #1 if the topic is "A," or to the broker #2 if the topic is "B."

Figure 11 shows a result from the packet monitoring on the server-side of the top broker, and Figure 12 shows a result from the packet monitoring on the broker #1 and the broker #2 side of the top broker. From Figure 11, we confirm that "publish" messages, addressed to the sensor #1 and the sensor #2, are transferred from the server (IP address: 192.168.1.2) to the top broker (IP address: 192.168.1.1). In addition, from Figure 12, we confirm that the "publish" message with topic "A", addressed to the sensor #1, is transferred only to the broker #1 (IP address: 192.168.2.2), and the message addressed to the sensor #2, "publish" message with topic "B" is transferred only to the broker #2 (IP address: 192.168.2.3). From this experimental result, we confirm that the top broker automatically transfers messages to each broker based on the topics of the messages published by the server.

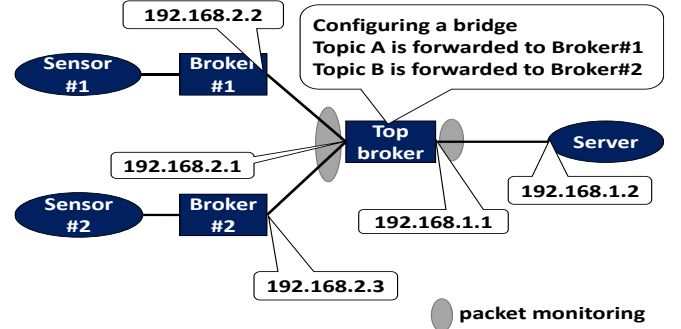


Figure 10: Experimental environment.

192.168.1.2	192.168.1.1	TCP	74 50236 → 1883 [SYN] Seq=0 Win=0
192.168.1.1	192.168.1.2	TCP	74 1883 → 50236 [SYN, ACK] Seq=0
192.168.1.2	192.168.1.1	TCP	66 50236 → 1883 [ACK] Seq=1 Ack=1
192.168.1.2	192.168.1.1	MQTT	103 Connect Command
192.168.1.1	192.168.1.2	TCP	66 1883 → 50236 [ACK] Seq=1 Ack=3
192.168.1.1	192.168.1.2	MQTT	70 Connect Ack
192.168.1.2	192.168.1.1	TCP	66 50236 → 1883 [ACK] Seq=38 Ack=5
192.168.1.2	192.168.1.1	MQTT	77 Publish Message [A]
192.168.1.2	192.168.1.1	MQTT	68 Disconnect Req
192.168.1.1	192.168.1.2	TCP	66 1883 → 50236 [ACK] Seq=5 Ack=52
192.168.1.1	192.168.1.2	TCP	66 1883 → 50236 [FIN, ACK] Seq=5 A
192.168.1.2	192.168.1.1	TCP	66 50236 → 1883 [ACK] Seq=52 Ack=6
192.168.1.2	192.168.1.1	TCP	74 50238 → 1883 [SYN] Seq=0 Win=0
192.168.1.1	192.168.1.2	TCP	74 1883 → 50238 [SYN, ACK] Seq=0 A
192.168.1.2	192.168.1.1	TCP	66 50238 → 1883 [ACK] Seq=1 Ack=1
192.168.1.2	192.168.1.1	MQTT	103 Connect Command
192.168.1.1	192.168.1.2	TCP	66 1883 → 50238 [ACK] Seq=1 Ack=3
192.168.1.1	192.168.1.2	MQTT	70 Connect Ack
192.168.1.2	192.168.1.1	TCP	66 50238 → 1883 [ACK] Seq=38 Ack=5
192.168.1.2	192.168.1.1	MQTT	77 Publish Message [B]
192.168.1.2	192.168.1.1	MQTT	68 Disconnect Req
192.168.1.1	192.168.1.2	TCP	66 1883 → 50238 [ACK] Seq=5 Ack=52
192.168.1.1	192.168.1.2	TCP	66 1883 → 50238 [FIN, ACK] Seq=5 A
192.168.1.2	192.168.1.1	TCP	66 50238 → 1883 [ACK] Seq=52 Ack=6

Figure 11: Server-side of the top broker.

192.168.2.1	192.168.2.3	MQTT	68 Ping Request
192.168.2.2	192.168.2.1	MQTT	68 Ping Response
192.168.2.1	192.168.2.2	TCP	66 60848 → 1883 [ACK] Seq=11 Ack=1
192.168.2.3	192.168.2.1	TCP	68 Ping Response
192.168.2.1	192.168.2.3	MQTT	66 35588 → 1883 [ACK] Seq=11 Ack=1
Raspberr_93:95:4d	Raspberr_93:8a:d3	ARP	60 Who has 192.168.2.1? Tell 192.1
Raspberr_93:8a:d3	Raspberr_93:95:4d	ARP	42 192.168.2.1 is at dc:a6:32:93:8
Raspberr_93:8a:d3	Raspberr_93:8a:d3	ARP	60 Who has 192.168.2.1? Tell 192.1
Raspberr_93:8a:d3	Raspberr_93:95:4d	ARP	42 192.168.2.1 is at dc:a6:32:93:8
Raspberr_93:8a:d3	Raspberr_93:95:4d	ARP	42 Who has 192.168.2.3? Tell 192.1
Raspberr_93:95:4d	Raspberr_93:8a:d3	ARP	60 192.168.2.3 is at dc:a6:32:93:9
192.168.2.1	192.168.2.2	MQTT	77 Publish Message [A]
192.168.2.2	192.168.2.1	TCP	66 1883 → 60848 [ACK] Seq=11 Ack=2
Raspberr_93:8a:d3	Raspberr_93:95:4d	ARP	42 Who has 192.168.2.2? Tell 192.1
Raspberr_93:95:4d	Raspberr_93:8a:d3	ARP	60 192.168.2.2 is at dc:a6:32:93:9
192.168.2.1	192.168.2.3	MQTT	77 Publish Message [B]
192.168.2.3	192.168.2.1	TCP	66 1883 → 35588 [ACK] Seq=11 Ack=2
Raspberr_93:8a:d3	Raspberr_93:95:4d	ARP	42 Who has 192.168.2.3? Tell 192.1
Raspberr_93:95:4d	Raspberr_93:8a:d3	ARP	60 192.168.2.3 is at dc:a6:32:93:9
192.168.2.1	192.168.2.2	MQTT	68 Ping Request
192.168.2.2	192.168.2.1	MQTT	68 Ping Response
192.168.2.3	192.168.2.1	TCP	66 1883 → 60848 [ACK] Seq=11 Ack=2
192.168.2.1	192.168.2.3	TCP	66 1883 → 35588 [ACK] Seq=11 Ack=2

Figure 12: Broker-side of the top broker.

## 6 CONCLUSION

In this paper, we propose a downward communication method for large-scale data collection systems such as a smart street lighting system. The proposed method consists of connection management by a server and downward transfer by tunneling. That is, in the proposed method, the server manages information of a broker, to which a sensor is connected, and transfers, by utilizing a tunneling technique, messages to the top broker that includes data to the sensor and information for the sensor managed by the server when the server notifies data to the sensor. Connection management by a server contributes to reducing the amount of memory required by the top broker. Firstly, to confirm the effectiveness of the proposed method in downward communication, we compared the communication traffic in the IoT core network of the IoT system for both the proposed method and a single broker approach, based on simplified analysis models. In the case of the IoT system utilizing a single broker, the amount of communication traffic increases rapidly when the number of sensors exceeds 100,000 due to the influence of memory overflow, however, the amount of communication traffic generated by the IoT system based on the proposed method increases linearly according to the number of sensors. Therefore, it is confirmed that the proposed method is effective even if the number of sensors is large. Secondly, to confirm the feasibility of the tunneling technique on the proposed method, we constructed a system consisting of multiple brokers and verified the function of transferring from a publisher to a subscriber using MQTT's bridge function by way of a specific broker being equivalent to a top broker. From this experimental result, we confirmed that the top broker automatically transfers messages to each broker based on the topics of the messages published by the server. To construct the proposed system in the future, it is necessary to create a prototype of the functions to be implemented in the top broker and to evaluate the communication traffic between the brokers in a real environment.

## ACKNOWLEDGMENT

A part of this work has been supported by the "Strategic International Standardization Promotion Project" of the Ministry of Economy, Trade, and Industry in Japan. The authors would like to heartily thank the members concerned.

## REFERENCES

- [1] Jie Ding, Mahyar Nemati, Chathurika Ranaweera, and Jinho Choi, "IoT Connectivity Technologies and Applications: A Survey," *IEEE Access* (Volume: 8) Apr. 2020.
- [2] Luís M. Borges, Fernando J. Velez, and António S. Lebres, "Survey on the Characterization and Classification of Wireless Sensor Network Applications," *IEEE Communications Surveys and Tutorials*, Vol. 16, No. 4, pp. 1860-1890. Apr. 2014.
- [3] Koichi Ishibashi, and Katsunori Yamaoka, "A Study of Network Stability on Wireless Sensor Networks," 9th International Conference on Next Generation Mobile Applications, Services and Technologies 2015, Sep. 2015.
- [4] Tetsuya Yokotani, and Yuya Sasaki, "Comparison with HTTP and MQTT on Required Network Resources for IoT," In 2016 international conference on control, electronics, renewable energy and communications (ICCEREC), pp. 1-6. Sep. 2016.
- [5] Yuya Sasaki, Tetsuya Yokotani, and Hiroaki Mukai, "MQTT over VLAN for Reduction of Overhead on Information Discovery." International Conference on Information Networking (ICOIN), pp. 354-356. May. 2019.
- [6] Koichi Ishibashi, Fuga Nakai, and Tetsuya Yokotani, "A Study of Data Collection Method by Using a Wildcard on a Large-Scale Smart Street Lighting System," The 24th World Multi-Conference on Systemics, Cybernetics and Informatics: WMSCI 2020, Sep. 2020.
- [7] Mohd. Saifuzzaman, Nazmun Nessa Moon, and Fernaz Narin Nur, "IoT Based Street Lighting And Traffic Management System," 2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC), Dec. 2017.
- [8] SCALAGENT "Benchmark of MQTT servers ActiveMQ 5.10.0 Apollo 1.7 JoramMQ 1.1.3 (based on Joram 5.9.1) Mosquitto 1.3.5 RabbitMQ 3.4.2," Jan. 2015.
- [9] Biswajeeban Mishra, "Performance Evaluation of MQTT Broker Servers," Lecture Notes in Computer Science International Conference on Computational Science and Its Applications -ICCSA, pp. 599-609. Jul. 2018.
- [10] Edoardo Longo, Alessandro E.C. Redondi, Matteo Cesana, Andres Arcia-Moret, and Pietro Manzoni, "MQTT-ST: a Spanning Tree Protocol for Distributed MQTT Brokers" 2020 ICC IEEE International Conference on Communications (ICC), Jun. 2020.
- [11] Pongnapat Jutadhamakorn, Tinnapat Pillavas, Vasaka Visoottiviseth, Ryousei Takano, Jason Haga, and Dylan Kobayashi, "A Scalable and Low-Cost MQTT Broker Clustering System," 2nd International Conference on Information Technology (INCIT) 2017, pp. 1-5. Nov. 2017.
- [12] Ryohei Banno, Jingyu Sun, Susumu Takeuchi, and Kazuyuki Shudo, "Interworking Layer of Distributed MQTT Brokers," *IEICE Transactions on Information and Systems*, Vol. E102.D, No. 12, pp. 2281-2294. Dec. 2019.
- [13] Eclipse Foundation, "Mosquitto™ An Open Source MQTT Broker," <https://mosquitto.org/>, accessed 2020-12-10.

Session 4:

Network

( Chair: Takuya Yoshihiro )





# A New Interest Forwarding Method Coping with The Publisher Migration in NDN

Taichi Iwamoto\* and Tetsuya Shigeyasu \*

\* Grad. School of Comprehensive Research, Pref. Univ. of Hiroshima, Japan  
 { r122001cf, sigeyasu }@ed.pu-hiroshima.ac.jp

**Abstract** – Recently, NDN (Named Data Networking) attracts a lot of attentions of network researchers as a network architecture based on content centric manner. NDN delivers requested contents by users according to the shortest path recorded in FIB (Forward Information Base) of CRs (Content Router). Incidentally, development of portable IT devices enables mobile publisher which generates contents under the migration environment. The current version of NDN, however, does not consider the publisher migration. So, NDN can not update the relay information in FIB even if a mobile publisher migrates its location. Wrong FIB information misleads Interest for new contents request from users and increase network traffics needlessly. In this paper, we propose a new Interest forwarding method to deal with the problem. The proposal construct adequate FIB information for mobile publisher location after the migration by sending empty DATA packet with R flag to the users. CRs on the delivered path of the DATA updates its FIB information when the reception of the DATA with R flag. The proposed method will be evaluated under the situations that mobile publisher changes its location during contents generations. The results of the evaluations confirm that our proposal well improve the performance of contents acquisition ratios.

## 1 INTRODUCTION

Recently, NDN (Named Data Networking)[1] attracts a lot of attentions of network researchers as a network architecture based on content centric manner. NDN delivers requested contents by users according to the shortest path recorded in FIB (Forward Information Base) of CRs (Content Router). Incidentally, development of portable IT devices enables mobile publisher which generates contents under the migration environment. The current version of NDN, however, does not consider the publisher migration. So, NDN can not update the relay information in FIB even if a mobile publisher migrates its location. Wrong FIB information misleads Interest for new contents request from users and increase network traffics needlessly.

In order to solve the problem, we have proposed the new Interest forwarding method which is the combination of content pre forwarding and erasing wrong FIB information, in the literature [2]. The literature have also reported that the proposal achieves higher content acquisition rate for transmitted Interest to the old location of mobile publisher, than traditional NDN.

On the other hand, contents created on mobile publisher after its migration, can be obtained only at the new location of the mobile publisher. Therefore, development of new method to disseminate new FIB information with regard to new location of mobile publisher, is strongly expected for NDN.

In this paper, we propose a new Interest forwarding method to deal with the problem. The proposal constructs adequate FIB information for leading Interests to new mobile publisher location information after the migration by sending empty DATA packet with R flag to the users. CRs on the delivered path of the DATA updates its FIB information when the reception of the DATA with R flag. The proposed method will be evaluated under the situations that mobile publisher changes its location during contents generations. The results of the evaluations confirm that our proposal well improve the performance of contents acquisition ratios.

## 2 RELATED WORKS

For the solution of the publisher migration problem on NDN, PMC (Publisher Mobility support protocol in CCN) has been proposed in [3]. In the PMC, in order to cope with the publisher migration, a publisher selects a *HomeNode*. The publisher registers its new location to the *HomeNode* when it migrates to the other place. By the *HomeNode*, Interests arrived at the old publisher's location due to the old/incorrect FIB entry, can be correctly forwarded to the new publisher's location. Fig.1 shows the procedure of Interest forwarding on PMC. As the figure shows, after a publisher migration, the publisher reports its destination information to the *HomeNode* as MR (Mobility Report) request. The *HomeNode* returns a MR response to the mobile publisher. At this time, nodes receiving MR response update its FIB entry according to the MR.

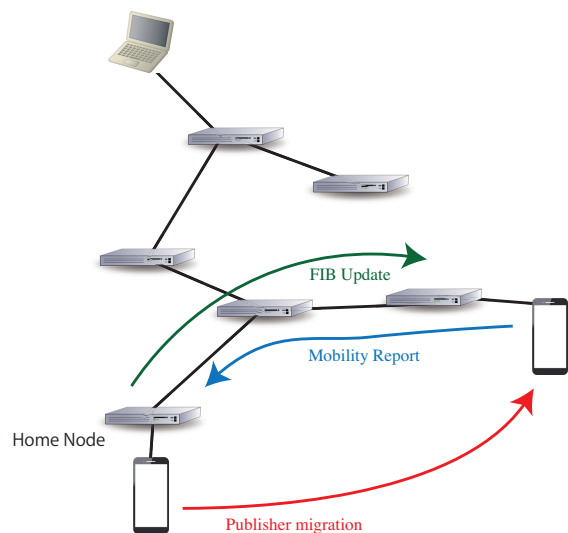


Fig.1 PMC.

### 3 PROBLEMS ON CONVENTIONAL NDN

This chapter mentions the three problems arisen at conventional NDN, induced by publisher migration.

#### 3.1 A problem losing access to the contents during migration of publisher

Once the mobile publisher starts to migrate its location to the other place, users who want to get the contents generated at the publisher, lose access to them. This is a losing access problem due to FIB information to the publisher become useless by publisher location change.

Fig.2 shows this problem. As shown in this figure, after the beginning of the migration of Host\_1, Interests destined to the Host\_1 can not reach the desired contents if the requested contents are not stored on CSs on CR1/CR2.

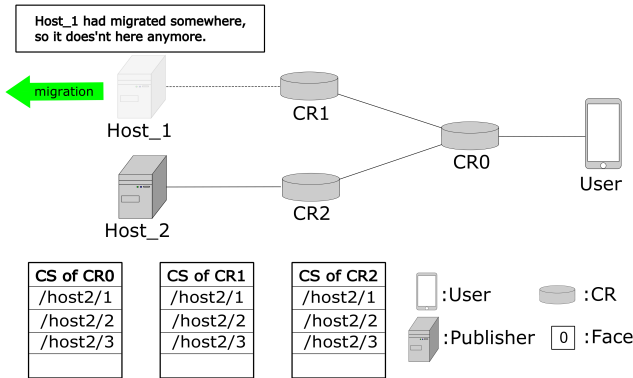


Fig.2 Problem on conventional NDN: losing access to contents during publisher migration.

#### 3.2 A problem losing access to the contents after migration of publisher

Old FIB information will be still holds on conventional NDN even if the publisher finished its migration. So, the Interests destined to the content generated by the publisher will be delivered to the old location. As shown in the Fig.3, Interest sent by User will be forwarded to the direction 1. This problem increases the unneeded network traffic.

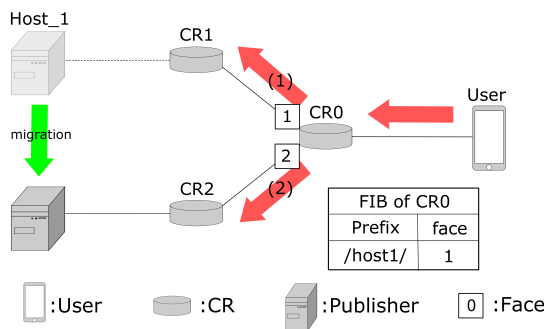


Fig.3 Problem on conventional NDN: Increasing network traffic due to wrong Interests delivering.

#### 3.3 A problem losing access to the contents after migration of publisher

As an optional procedure, multicast forwarding of Interest method is implemented to increase content acquisition rate on conventional NDN. It would be thought that the NDN turns on this multicast forwarding method to increase contents acquisition rate when the content acquisition rate becomes low due to the publisher migration. However, this situation also increases both of network traffic due to the Interest flooding, and PIT entries on CRs locating out side shortest-path to the new publisher location (Fig.4).

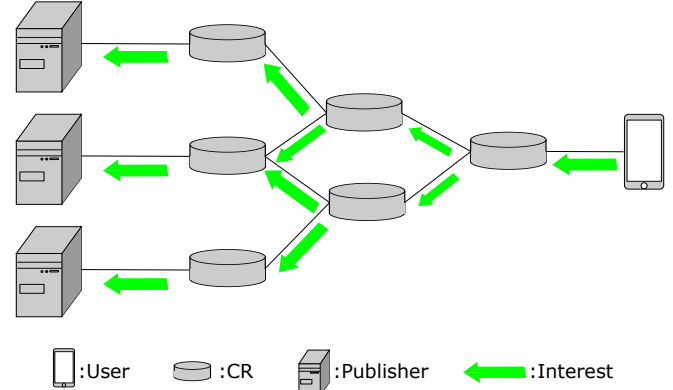


Fig.4 Problem on conventional NDN: increasing amount of multicast traffic due to losing access to contents after publisher migration.

### 4 PROPOSAL

In this chapter, we propose a new solution for coping with the problem induced by the publisher migration on conventional NDN. Our solution consists two parts: 1) contents pre-forwarding and deletion of old (useless) FIB information, and 2) dissemination to notify new FIB information.

Although our solution is a hybrid schema of the above two parts, first part is same procedure as the our past proposal in the literature [2], and the second part is a new procedure first time appeared in this paper.

#### 4.1 A method for contents pre-forwarding and deletion of useless FIB information

As we mentioned in the previously, during the publisher migration, all users willing to get contents lose access to the contents if the desired contents are not stored at the intermediate CRs' CS. Then, this section proposes the contents pre-forwarding to the neighboring CR before publisher migration. By the pre-forwarding, Interest can be delivered to the CR holding desired contents even when the publisher migration. This would increase content acquisition rate. By utilizing the pre-forwarding, we also propose to delete the old (wrong) FIB information on CRs belonging to the shortest path to the old publisher location. For realizing the deletion procedure, content caches delivered by pre-forwarding marked with *m-flag* (Migration flag). The *m-flag* indicates that the publisher migration. CRs received the

contents with the *m-flag* deletes old FIB information corresponding to the content name.

Fig.5 shows the procedure of this mechanism.

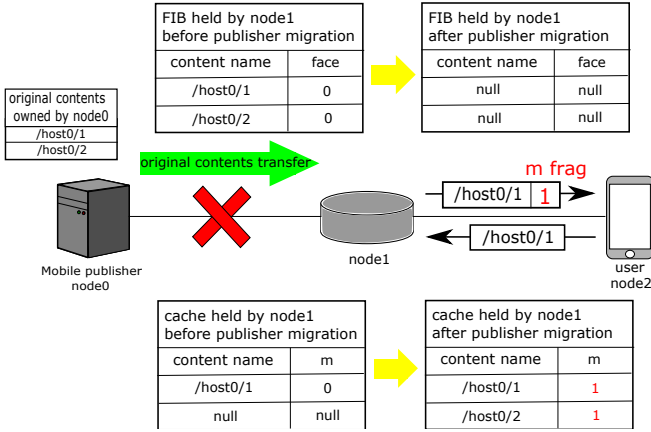


Fig.5 Procedure of a method for contents pre-forwarding and deletion of useless FIB information.

Detailed procedure are described as below:

1. Procedure on CR when receiving contents with *m-flag*
  - A) Delete its FIB information corresponding to the same content name with the received content.
  - B) Store the received content into its CS with *m-flag*.
  - C) Forward the received content with *m-flag* if the corresponding entry is recorded into its PIT.
2. Procedure on CR when receiving the contents without *m-flag*.
  - A) Search FIB information having same content name with the newly received content. Register the new FIB with the information of received face if the corresponding FIB information is not recorded yet.
  - B) Search CS whether content with same content name is cached or not. If no, store the received content. When the content having same content name is already cached, CR further checks the *m-flag* is set. If the content has the *m-flag*, unset the *m-flag*.
  - C) Forward the received content without *m-flag* if the corresponding entry is recorded into its PIT.
3. Procedure on CR when it receives the Interest (same procedure with the conventional)
 

CR returns corresponding content when it holds desired cache. If the CR does not hold the requested content, it record the requested information on its PIT, and forward the Interest to the its upstream CR.

## 4.2 A method for dissemination of new FIB information

In this section, we propose a second method which renew the FIB information on CRs to realize fast content delivery to users and reduce increasing unneeded traffic.

For the purpose, we use *R-flag* (*Rebuild flag*) in addition to the current Interest header. By referring the state of *R-flag*,

CR belonging to the shortest path of new content delivery can quickly renew its FIB information.

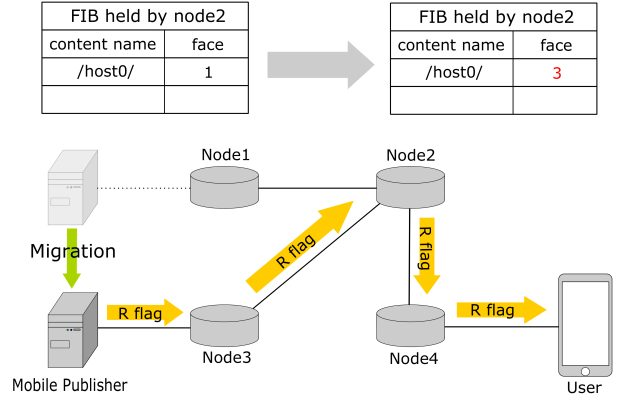


Fig.6 Procedure of dissemination of new FIB information.

Fig.6 shows the procedure using *R-flag*. In this method, mobile publisher transmits new content packet with *R-flag* when it finished its migration. CRs receiving the content with *R-flag*, set new FIB information according to the arrival face of the content with *R-flag*. By this procedure, CRs can obtain desired content without turning on the multicast Interest forwarding which increases needless traffic. The content with *R-flag* will be forwarded to the downstream CRs until end users.

The detailed procedure introducing the *R-flag* is described as below:

1. Procedure on CRs receiving content with *R-flag*
  - A) Check the FIB information relating to the same content name with the content with *R-flag*. If it does not exist, the CR newly adds the FIB information.
  - B) Store the received content into its CS
  - C) Forward the content with *R-flag*, to the downstream CRs. For reducing the forwarding traffic, the content will only be forwarded to the downstream CRs which received the similar content name.
2. Procedure on Users receiving content with *R-flag*

A user receiving content with *R-flag*, checks its FIB information. If there is no corresponding FIB entry, the user newly adds the FIB entry according to the received content with *R-flag*.

## 5 PERFORMANCE EVALUATION

For clarifying the availability of the proposal, this chapter evaluates the performance by computer simulation.

### 5.1 Evaluation environment

This section describes the evaluation environment. Table.1 shows the parameters used by simulation.

Table.1 Simulation parameters

Parameter	Value
Number of nodes	24
Interest generation rate	100 [pkt/sec]
Number of contents	1,000
Interest Packet size	1,024 [bytes]
Content Packet size	1,024 [bytes]
Cache capacity	infinity

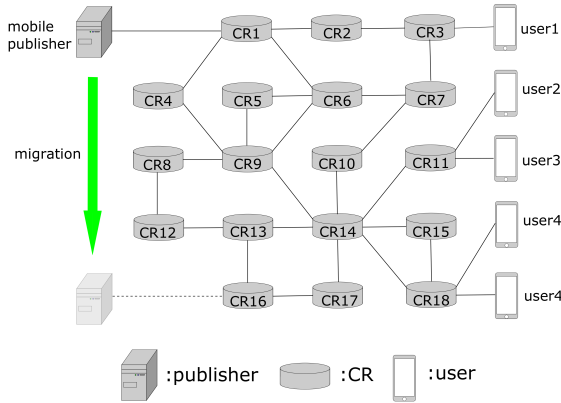


Fig.7 Simulation topology.

Fig. 7 shows the simulation topology. As shown in the figure, number of publisher, CRs and users are, 1, 18, 5, respectively. Mobile publisher starts to migrate from the neighbor of CR1 to the neighbor of CR16, during the simulation. Content requests will be arrived at all users with rate 100 [request/sec].

Total number of original content published by the mobile publisher is 1,000. Then, content name recorded in Interest is randomly selected with the range. Pre-forwarding contents are also randomly selected at the beginning of publisher migration.

On the simulation, amount of pre-forwarding content, TTL (Time To Live) of Interest, and capacity of PIT are varied for clarifying the characteristics of the proposal. For the evaluations, content acquisition rate is derived as a value of number of contents obtained divided by number of contents requested.

## 5.2 A relationship between content acquisition rate and amount of pre-forwarding content

This section reports the characteristics of content acquisition rate under varying amount of pre-forwarding content. Values of TTL and PIT are 0.1 [s] and infinity, respectively.

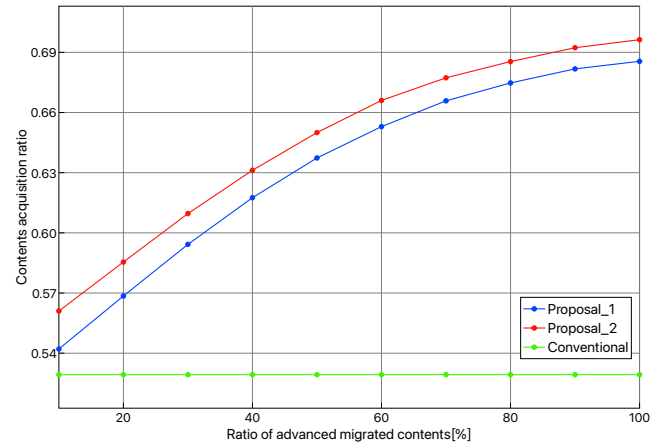


Fig.8 Characteristics of content acquisition rate - amount of pre-forwarding content.

Fig.8 shows the results of computer simulation. In this figure, Proposal\_1 and Proposal\_2 show the characteristics of methods implementing only pre-forwarding, and both of pre-forwarding and *R-flag*. In addition, Conventional means the performance of original NDN.

We can see from the figure, Proposal\_1 and Proposal\_2 increase content acquisition rate accordance with the amount of pre forwarding content. In addition, Proposal\_2 always higher performance than Proposal\_1. The difference among two methods are thanks to fast re-build of Interest forwarding path by *R-flag*.

## 5.3 A relationship between content acquisition rate and length of TTL

This section reports characteristics of content acquisition rate under varying length of TTL. For this evaluation, all contents are forwarded to neighbor of publisher at the beginning of publisher migration.

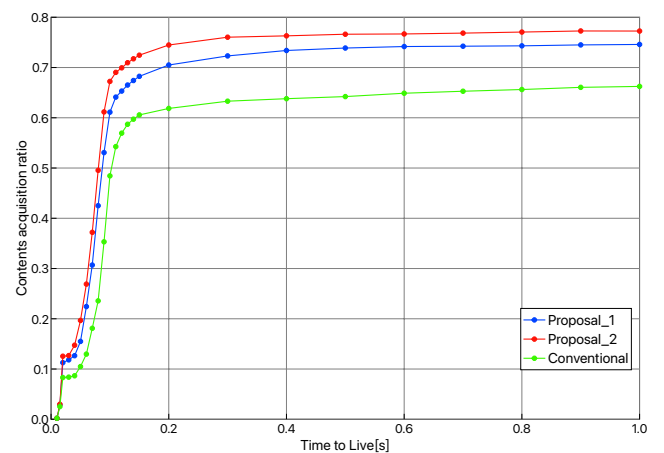


Fig.9 Characteristics of content acquisition rate - amount of pre-forwarding content.

Fig.9 shows the results of the evaluation. In this figure, horizontal axis shows the length of TTL, and the vertical axis shows content acquisition rate. Colored lines shows same mean of the previous figure.

It is obviously that, all methods increase the content acquisition rate accordance with increasing the length of TTL. In addition, as the figure shows, increasement of content acquisition rate becomes small over the TTL is larger than 0.1 [s]. This is because the average RTT (Round Trip Time) for content acquisition are 0.2 [s]. Then, if we enlarged the TTL larger than the RTT, It does not contribute to increase content acquisition rate.

As the figure shows, Proposal\_2 always keeps the highest performance than the two methods.

#### 5.4 A relationship between content acquisition rate and PIT capacity

This section evaluates the characteristics of content acquisition rate and PIT capacity. We use the 0.1 [s] as the TTL on the evaluation.

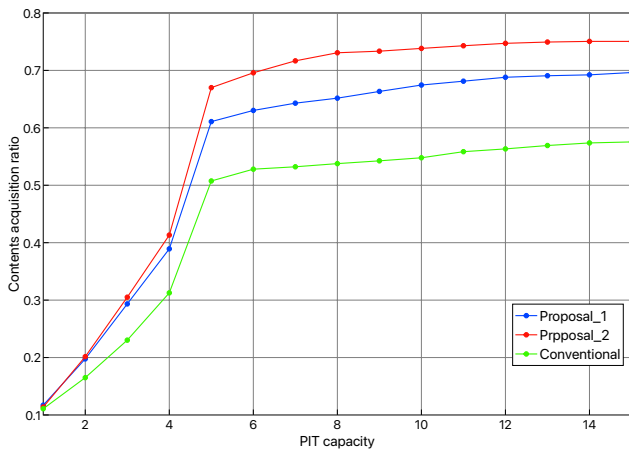


Fig.10 Characteristics of content acquisition rate – PIT capacity.

Fig.10 shows the results of the evaluation. As the figure shows, content acquisition rate increases accordance with the PIT capacity under the PIT capacity is smaller than 5. However, increasements of content acquisition rate becomes small when the PIT capacity is larger than 5.

As we can see this figure, Proposal\_2 always keeps advantage than the other two methods.

## 6 CONCLUSION

As we know, conventional NDN is developed to reduce traffic for content deliveries. NDN only assumes that content publisher is fixed, then, Interest forwarding is also fixed.

However, the developments of ICT technologies, mobile publisher generating contents during its migration, is strongly expected. To cope with the appearance of such mobile publisher, NDN must realize appropriate Interest/content forwarding while dealing with such publisher migration.

In this paper, we discussed the problems of publisher migration on conventional NDN. To cope with the problem, we proposed *m-flag* and *R-flag* procedure.

*m-flag* procedure is destined for deletion of wrong (old) FIB information to reduce unneeded Interest forwarding. *R-flag* procedure is destined for fast establishing of new Interest appropriate forwarding.

The results of computer simulations confirm that our proposal well improves the content acquisition rate under the situation that publisher migrates its location during the simulation.

However, our proposal consumes cache buffer of CS on CR of neighbor of publisher. Off course, this procedure is necessary for improve the content acquisition rate during publisher migration. The pre forwarded contents, however, becomes unneeded buffer consumption after the new FIB information is disseminated after mobile publisher finished its migration and informs it by transmission of content with *m-flag*.

For the future, we would like to deal with this unneeded buffer consumption on our proposal.

## REFERENCES

- [1] Named Data Networking (NDN): A Future Internet Architecture (online), <https://named-data.net> (accessed 2021-6-1).
- [2] T. Iwamoto and T. Shigeyasu, A Study on Reducing Interest Misleading by Publisher Migration on Mobile Networks, Proc. of BWCCA2020, Communication and Applications, pp.407-415, 2020.
- [3] Han, M. Lee, K. Cho, T. Kwon and Y. Choi, "Publisher mobility support in content centric network," of Proc. of IEEE International Conference on Information Networking (ICOIN), pp.214-219, 2014.





# A Proposal on mechanisms of ICN with traffic control functions for IoT communication

Atsuko Yokotani\*, Hiroshi Mineno\*, Satoshi Ohzahata\*\*, Tetsuya Yokotani\*\*\*

\*Graduate School of Informatics, Shizuoka University, Japan

{yokotani.atsuko20@, mineno@inf.}shizuoka.ac.jp

\*\*Graduate School of Informatics and Engineering, University of Electro-Communications  
ohzahata@is.uec.ac.jp

\*\*\*College of Engineering, Kanazawa Institute of Technology  
[yokotani@neptune.kanazawa-it.ac.jp](mailto:yokotani@neptune.kanazawa-it.ac.jp)

**Abstract** – Information Centric Network (ICN) is a promising candidate to mitigate protocol overheads to transfer information. Therefore, although it can be concluded that ICN can be applied to the Internet of Things (IoT), traffic control functions in ICN have not been specified to real-time services in IoT. This paper proposes detailed mechanisms on bandwidth reservation in traffic control functions, and describes their performance evaluation by network traffic emulation.

**Keywords:** IoT, ICN, Traffic control, Real-time application

## 1 INTRODUCTION

The Internet of Things (IoT) is a worldwide topic of interests. As various services utilizing IoT are deployed, the communication network plays an important role. Most IoT services expect wide-area network services, including Internet services [1]. However, in the mature stage of IoT services, if these services are deployed over the Internet as it currently exists, some serious problems will be highlighted, e.g., large overheads of the legacy protocols, processing resources of their overhead in communication equipment, and processing power of Internet protocol (IP) address translation by the domain name system (DNS).

To mitigate these problems, Information-Centric Network (ICN) technologies have been discussed to facilitate IoT services. ICN technologies invoke independent communication of IP. They also provide networked cache to reduce duplicate traffic transfer.

This paper proposes communication sequences for IoT services based on ICN. It also proposes operations with traffic control mechanisms with prioritized traffic flows on ICN base networks for low latency IoT services. Although ICN technologies include various options, this paper focuses on Content Centric Network (CCN) [2] which is the most popular mechanism.

## 2 SUMMARY OF CCN OPERATIONS

CCN can provide simplified communication sequences to obtain information from servers. Figure 1 shows the concept in CCN. In CCN, a user accesses the server using content names directly, without translation between an IP address and URL indicating the location of the content. Moreover,

the transferred content can be temporarily stored at some intervening, or interworking, points in networks. When another user would access the server to obtain that content, the interworking point (IWP) provides the requested content from its cache instead of the server. Therefore, duplicate transfer of contents by the server and the limited processing power of content transfer in the server are mitigated. In CCN, a request for content and the response including the content are referred to as **Interest** and **Data** messages, respectively.

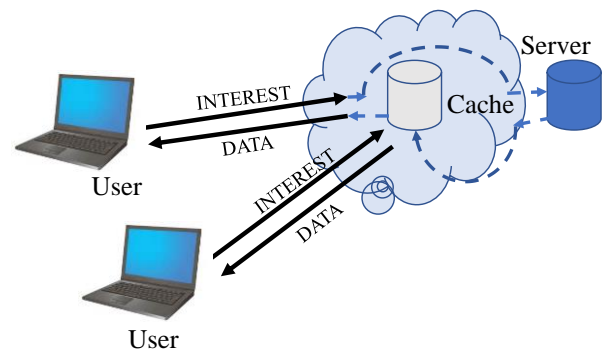


Figure 1 Concept in CCN

In CCN, there are three components to an IWP (i.e., Content Store (CS), Pending Interest Table (PIT), and Forwarding Information Base (FIB)), which handle **Interest** and **Data**. The relationship among these components is shown in Figure 2. IWPs are connecting points between CCN links and can be positioned as routers in the current Internet.

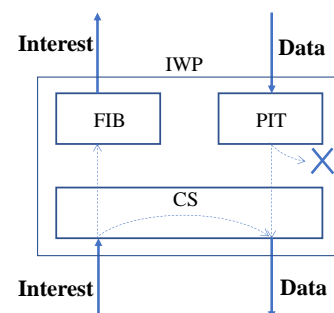


Figure 2 Operation among components in an IWP

### 3 COMMUNICATION SEQUENCES IN IOT SERVICES

To deploy IoT services, communication sequences are classified into three types, as shown in Figure 3 [3]. Generally, an end device includes various sensors, actuators, and a communication interface accommodating them. Topology and communication protocols between sensors, actuators, and the communication interface have various options. This system is relatively small and is configured by dedicated networks, so it can be optimized for specific services.

In these types, Type 1 seems to be in the majority because most IoT services require information from a large number of end devices, including sensors as end points of communication. In these sequences, some interworking points relay information of IoT services.

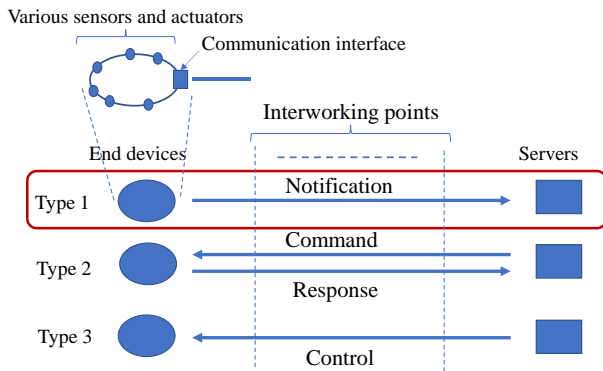


Figure 3 Types for IoT services in communication sequences

When CCN is applied to IoT services, it provides simpler communication for IoT services than the conventional Internet. For example, in IoT services, the huge number of end devices create tiny information blocks and transfer these blocks across networks, e.g., Type 1 in Figure 3. In this situation, large protocol headers, i.e., the hypertext transfer protocol (HTTP), and three-way handshake procedures in transmission control protocol and IP (TCP/IP) causes an increase in traffic volume and processing power. Moreover, processing in DNS generates a heavy load for communication equipment.

CCN is designed to mitigate these problems in deployment of IoT service. However, it may cause a security issue. Especially in the case of Type 1 of Figure 3, when suspicious devices are connected to networks, distributed denial of service (DDoS) attacks may be initiated. This problem has been indicated in [4]. In that paper, authors proposed that an interworking point in networks could provide a screen of transfer traffic prior to endpoints as one of regulation mechanisms on incoming IoT traffic, as shown in Figure 4 [4].

Generally, in the case of Type 1, end devices transfer information periodically to networks. In particular, real-time services in IoT will require periodic transfer sequences [5]. Requirements of these services are surveyed in [5] as described in Table 1.

In Figure 4 as an example, if data transfer is performed by less than the half cycle period identified in Table 1, the cycle time in Table 1 is guaranteed.

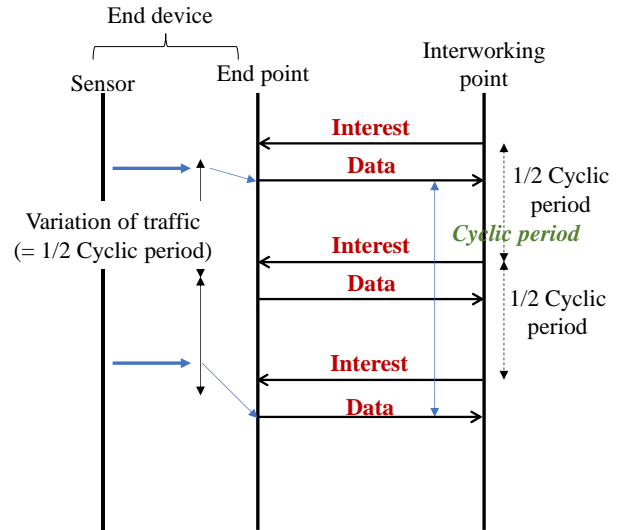


Figure 4 Cyclic transfer in Type 1

Table 1 Summary of traffic requirements of low latency IoT services

	Latency (ms)	Packet loss ratio	Cycle (ms)	Size (B)	Device density
Factory	0.25~10	$10^{-9}$	0.5~50	10~500	0.33~3/m <sup>3</sup>
Industrial plant	50~100	$10^{-4}$ ~ $10^{-3}$	100~5000	40~100	10000/Plant
Smart grid	3~20	$10^{-6}$	10~100	80~1000	10~2000/km <sup>2</sup>
Transportation (Safety drive)	10~100	$10^{-5}$ ~ $10^{-3}$	100~1000	~1000	500~3000/km <sup>2</sup>

### 4 PROPOSED MECHANISMS OF CCN FOR IOT SERVICES

In Section 3, the possibilities and new issues of IoT services based on CCN were described. However, CCN for IoT services have not provided traffic control functions. In this section, basic mechanisms based on the concept for IoT services by CCN are proposed. Then, mechanisms with traffic control functions focusing on reservation of bandwidth are proposed.

#### 4.1 Basic mechanisms for IoT services

Basic transfer mechanisms using CCN technologies are shown in Figure 5.

In Figure 5, end devices are deployed according to offered services, e.g., monitoring of industry plants. At first, Server #1 submits an **Interest** to obtain information through some Interworking points (IWP). Then, end devices reply to Server #1 with **Data** containing target information. After that, IWP #1 transfers that **Interest** periodically according to provisioning timing, e.g., required cycles. As the PIT in each IPW is set after the first **Interest**, **Data** is transferred to each IPW and can be stored in cache (Content Store) for every periodic **Interest**.



If other servers, e.g., Server #2, intend to obtain information regarding the **Interest**, IWP#3 updates the PIT to indicate requested information by a new server and then provides the **Data** stored in cache.

The processing sequences of **Interest** at IWP #1 shown in Figure 5 should be modified as in Figure 6. The processing sequences of **Data** at IWP #2 should be added to the original sequences shown in Figure 2. Other sequences are compiled in the original sequences shown in Figure 2.

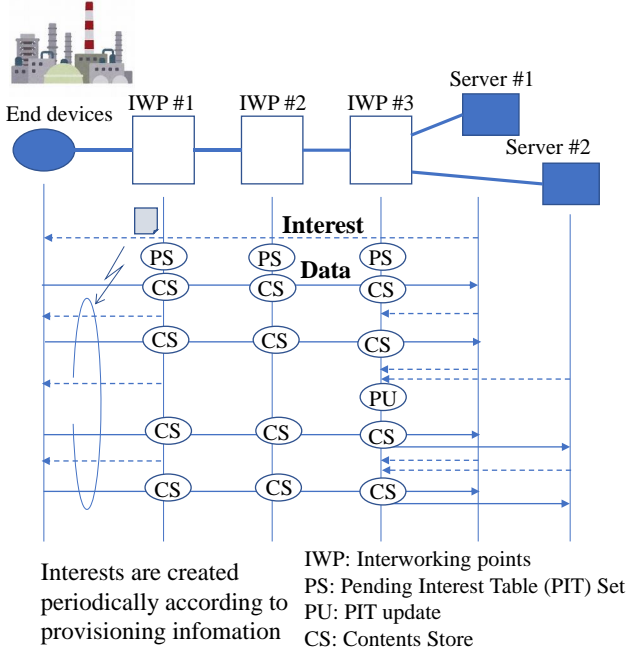


Figure 5 Operations in basic transfer mechanisms

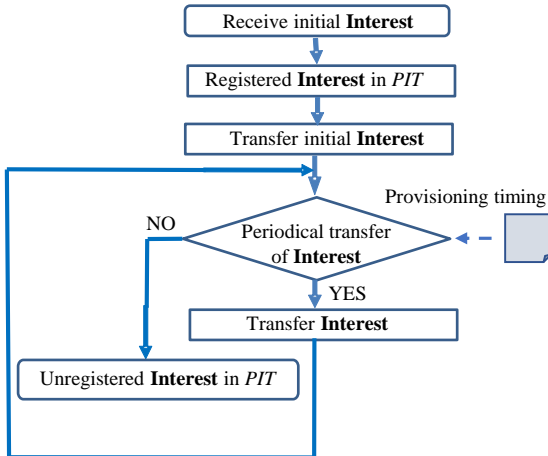


Figure 6 Detailed operations in proposed mechanisms

## 4.2 Mechanisms with traffic control

In Section 4.1, basic transfer mechanisms were proposed. However, when services are aggregated on networks, traffic control functions should be provided to prioritize IoT services requiring low latency. For this purpose, bandwidth reservation of IWPs are proposed.

A summary of these proposed mechanisms is shown in Figure 7.

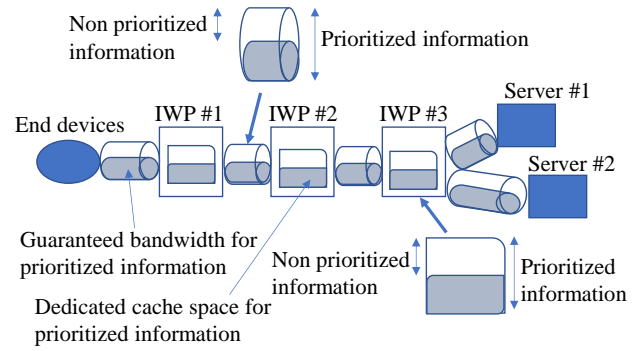


Figure 7 Traffic control functions in networks

In Figure 7 each link has guaranteed bandwidth for prioritized information, which consists of an **Interest** and **Data** pair. In this system, prioritized information can utilize the full capacity if non-prioritized information is not transferred. In each IWP, dedicated cache space is assigned for prioritized **Data**. Full cache space can be assigned for prioritized information if non-prioritized information is not stored in cache.

With these traffic functions guaranteed bandwidth is reserved by the dual leaky bucket mechanism [6]. The bandwidth less than the commitment information rate (CIR) should be reserved for prioritized information. The bandwidth of more than the CIR and less than the sustainable information rate (SIR), e.g., link rate, can be reserved for prioritized information according to availability of non-prioritized information.

## 5 PERFORMANCE EVALUATION

In this section, the proposed mechanisms are evaluated using the CCN software platform referred to as CCNx [7].

### 5.1 Network configuration

The network configuration for performance evaluation is shown in Figure 8. In Figure 8, an IWP connects end devices and eight servers. The two servers processed prioritized information. Other servers processed non-prioritized information. Therefore, in the link between end devices and the IWP, bandwidth was reserved for prioritized information.

In this performance evaluation, end devices, servers, and IPWs were emulated using Linux PCs with CCNx.

### 5.2 Numerical examples

In this performance evaluation, bandwidth was normalized because CPU in PCs cannot emulate a real bit rate. Parameters from the performance evaluation were as follows:

- Link rate (=SIR) 10 M/unit
- Reserved bandwidth (=CIR) 3 M/unit
- Information block size 4 kB
- Generating rate of Interest for Prioritized information 50/unit

- Generating rate of Interest for non-Prioritized information 50/unit
- Dedicated space of cache 1500
- Total space of cache 4500
- Update cycle of dedicated space 1 unit
- Basic policy of cache LRU

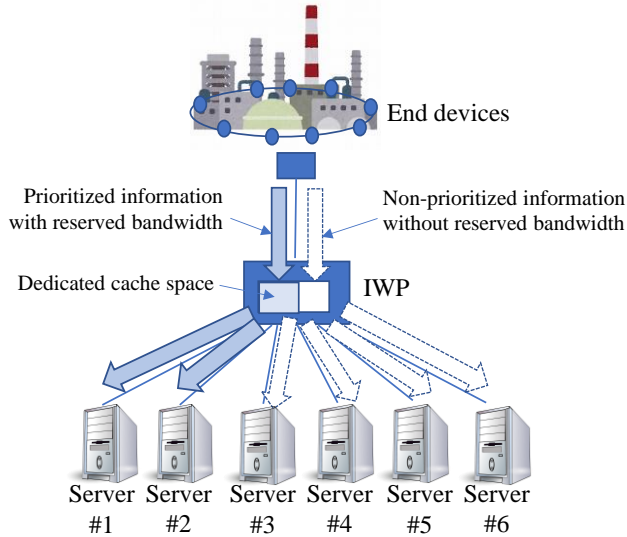


Figure 8 Network configuration for performance evaluation

### 5.3 Numerical examples

Some numerical results are shown in Figure 9.

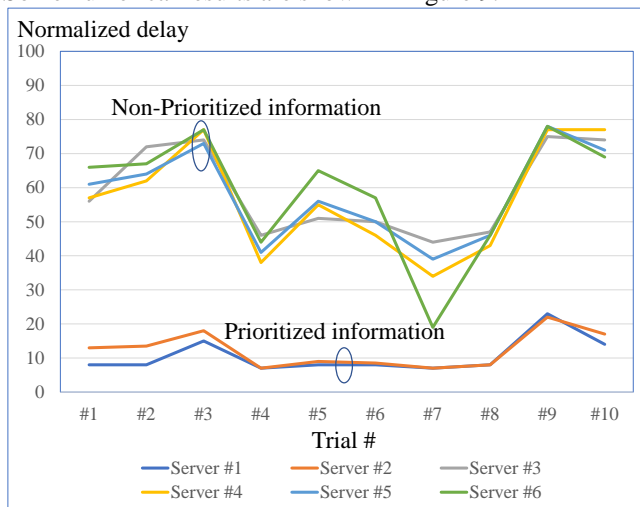


Figure 9 Example of observed results

This graph shows latency between **Interest** and **Data** in each server. The number of trials was 10 times, as shown in the horizontal axes. Vertical axes show the latency by relative values. It is indicated as “Normalized delay”.

In this case, bandwidth reservation for prioritized information was always activated. Moreover, when the two IWPs case is compared with the one IWP case, prioritized information with cache control is almost same. It can be concluded that delay of prioritized information is relative stable independent of system scale if cache control is activated.

## 6 CONCLUSIONS

This paper proposed transfer mechanisms in IoT communication using ICN technologies. ICN technologies can solve communication problems of IoT presented by the current Internet and provide some advantages in IoT communication. However, security issues, e.g., DDoS attack, must be considered. This paper focused on CCN as a typical mechanism, and proposed communication sequences based on CCN to solve this issue.

Moreover, to comply with low latency requirements of IoT services, we proposed traffic control functions, including bandwidth reservation and cache control. Finally, this paper confirmed the advantages of the proposed mechanisms by the emulated system.

In future works, authors will provide more detailed evaluation and will promote these proposals to the international standardization, e.g., ITU-T, ISO/IEC JTC1 and other fora.

## REFERENCES

- [1] T. Yokotani, and K. Kawai, “Concepts and requirements of IoT networks using IoT Data Exchange Platform toward International standards”, IEEE Conference on Standards for Communications and Networking (IEEE CSCN), #1570570960, 2019, DOI: 10.1109/CSCN.2019.8931337, IEEE Xplore
- [2] V. Jacobson, D. K. Smetters, J. D. Thornton, M. Plass, N. Briggs, and R. Braynard, “Networking Named Content,” ACM CoNEXT 2009, pp.1-12, 2009.
- [3] T. Yokotani, S. Yamamoto, S. Ohno, K. Sasabayashi, and K. Ishibashi, “Survey and comparison of Interworking point routing mechanisms for IoT services in wide area ICNs”, International Conference on Emerging Technologies for Communications (ICETC 2020), D2-4, 2020
- [4] A. Yokotani, H. Mineno, and T. Yokotani, “A proposal on the access control mechanism for real time IoT services using ICN technology”, IoT Enabling Sensing/Network/AI and Photonics Conference 2021 (IoT-SNAP 2021), IoT-SNAP-5-06, 2021
- [5] P. Schulz, M. Matthe, H. Klessig, M. Simsek, G. Fettweis, J. Ansari, S. Ali Ashraf, B. Almeroth, J Voigt, I. Riedel, A. Puschmann, A. Mitschele-Thiel, M. Muller, T. Elste, and M. Windisch, “Latency critical IoT applications in 5G: Perspective on the design of radio interface and network architecture”, pp. 70-78, IEEE Communication Magazine, February 2017
- [6] Atsuko Yokotani, Satoshi Ohzahata, Ryo Yamamoto, and Toshihiko Kato, “A Dynamic Cache Size Assignment Method with Bandwidth Reservation for CCN”, 2019 International Conference on Information Networking (ICOIN 2019), P2-15, 2019, DOI: 10.1109/ICOIN.2019.8718174, IEEE Xplore
- [7] “CCNx project”, <http://www.ccnx.org>

# A Study on Assistant Devices for Presentation of Distinctive Viewers' POV in 360-degree Internet Live Broadcasting

Masaya Takada\* and Yoshia Saito\*

\*Graduate School of Software and Information Science, Iwate Prefectural University, Japan  
m-osawada@iwate-jh.ed.jp, y-saito@iwate-pu.ac.jp

**Abstract** - In this paper, we propose three assistant devices for presentation of the distinctive viewers' POV in 360-degree Internet live broadcasting. The first one is a belt-type device which presents directions of the POV using vibration motors. The second one is a LED-type device which presents directions of the POV using LED light sources cylindrically. The third one is a robot-type device which presents directions of the POV using movement of the robot's head. We developed the three assistant devices and conducted an evaluation experiment. From the experiment, we found the broadcasters did not prefer assistant devices in visual form and they had a positive impression of the robot-type device.

**Keywords:** 360-degree Internet live broadcasting, Viewers' POV, Assistant devices

## 1 INTRODUCTION

In recent years, many people use Internet live broadcasting services. In the Internet live broadcasting services, the viewers can enjoy real-time communication with the broadcaster. Besides, YouTube started a 360-degree Internet live broadcasting service which supports omnidirectional cameras from 2016. It enables anyone to easily use the 360-degree Internet live broadcasting service now.

In the 360-degree Internet live broadcasting, a broadcaster takes a 360-degree video using an omnidirectional camera and distributes it to viewers in real-time via the Internet. The broadcaster does not need to care about the view angle of the camera. The viewers can change their point of view (POV) while watching the 360-degree live video and they can watch the video from POV which they are interested in.

The 360-degree Internet live broadcasting, however, has an issue that the broadcaster cannot check the viewers' POV. In the conventional Internet live broadcasting, it uses a web camera which has a single lens and the single lens definitely shows the rectangular photographing range. The broadcaster can be aware of the viewers' viewing range and what they are watching by direction of the lens. On the other hand, in the 360-degree Internet live broadcasting, it uses an omnidirectional camera which has a wide-angle lens or multiple lens and the broadcaster cannot know what the viewers are watching by direction of the lens.

There are many studies about the role of gaze information in the remote communication [1][2]. In the studies, it is turned out that the communicatee's gaze information indicates the target of interest or center of the topic. The gaze information in the remote communication is similar to

the viewers' POV in the 360-degree Internet live broadcasting. Therefore, the viewers' POV are not only information which indicates where the viewers are watching but also information which indicates what the viewers are interested in. Because of that, the broadcaster sometimes cannot understand the context of the viewers' comments and it can be a factor which causes communication errors between the broadcaster and the viewers.

To solve this issue, we have studied about an algorithm which detects distinctive viewers' POV to grasp viewers' interests [3]. In this research, the algorithm could detect useful viewers' POV for the broadcaster. We have also studied the effect of presentation of the distinctive viewers' POV to the broadcaster [4]. In this research, we found the presentation of the distinctive viewers' POV could be effective for the broadcaster and it gave positive effects to the communication between the broadcaster and the viewers. It enabled the broadcaster to know what the viewers were interested in. The broadcaster also had a chance to communicate with passive viewers who sent few comments to the broadcaster by the presentation of the distinctive viewers' POV. Even if the distinctive viewers' POV which were not useful to know the viewers' interests were displayed, it did not have a significant negative impact on communication and broadcasting.

The use case which we envisioned for the previous researches were that a single broadcaster delivered the situation of walking through a tourist spot. The broadcaster would visit a tourist spot and report about the spot to the viewers. The equipment used for the broadcasting were a laptop computer and an omnidirectional camera. The broadcaster carried a backpack with a camera mounter to fix the omnidirectional camera and handed the laptop PC. The distinctive viewers' POV were simply presented as red circles on the equirectangular video using a laptop PC for the broadcaster. The presentation method was not easy to comprehend and the use of a laptop PC would increase the risk of accidents.

To solve these issues, we study several assistant devices to present the distinctive viewers' POV to the broadcaster in an effective manner. In this research, we propose three assistant devices for presentation of the distinctive viewers' POV in 360-degree Internet live broadcasting. The first one is a belt-type device which presents directions of the POV using vibration motors. The second one is a LED-type cylindrical device which presents directions of the POV using LED light sources. The third one is a robot-type device which presents directions of the POV using movement of the robot's head.

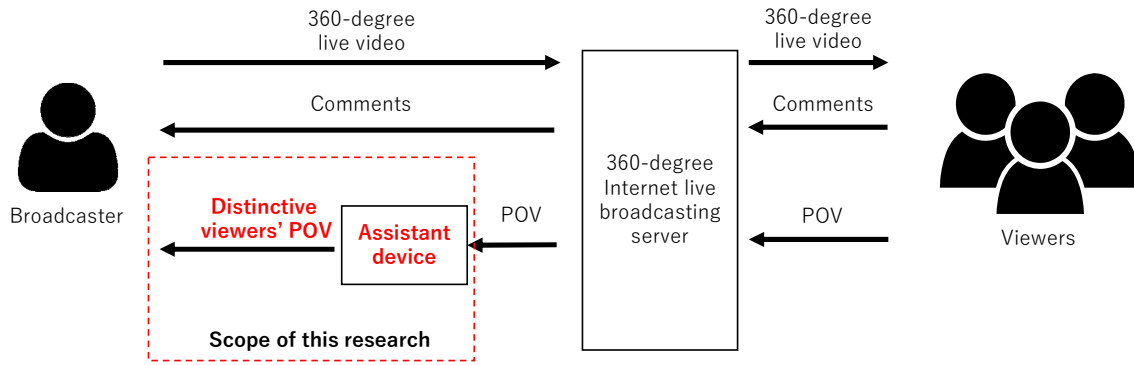


Figure 1: A model of the assistant devices in the 360-degree Internet live broadcasting

The contributions of this paper are summarized as follows:

- We developed three assistant devices for presentation of distinctive viewers' POV.
- We clarified the effects of each assistant device in experiments and which device the broadcaster preferred.

The rest of this paper is organized as follows. Section 2 describes three assistant devices for presentation of the distinctive viewers' POV in 360-degree Internet live broadcasting. Section 3 describes implementation of the assistant devices. Section 4 describes evaluation experiments to clarify the effects of each assistant device. Section 5 summarizes this study.

## 2 ASSISTANT DEVICES

### 2.1 A model of the assistant devices

We propose assistant devices for presentation of distinctive viewers' POV in 360-degree Internet live broadcasting. Figure 1 shows a model of the assistant devices. In the model, a broadcaster sends 360-degree live video to the broadcasting server and viewers can watch the broadcasting and send comments to the broadcaster accessing the broadcasting server as same as typical 360-degree Internet live broadcasting services. The viewers can watch the video changing their POV. The coordinates of the POV are also sent to the broadcasting server in real time as same as our previous work. In this research, the broadcasting server sends the POV to an assistant device. The assistant device uses the algorithm of the previous work to detect distinctive viewers' POV. The assistant device presents the distinctive viewers' POV to the broadcaster by using a presentation means. The broadcaster can communicate with the viewers smoothly estimating their interests from the presentation of the distinctive viewers' POV.

### 2.2 Requirements of the assistant devices

The broadcaster's motivation to use the assistant devices is to make the walk in the tourist spot better. The idea is that

the broadcaster can gain a companion from the viewers, even if the broadcaster is traveling alone. In this environment, there are the following three requirements.

1. It should be short time to check the assistant device.
2. It should not restrain the broadcaster's physical activity.
3. It should improve the broadcaster's experience.

The first and second requirements are needed not to increase the risk of accidents, for example, in case of using a laptop PC. The third requirement means the assistant device has great potential to become a companion for the broadcaster in order to improve his/her experience.

### 2.3 Ideas of the assistant devices

Based on the requirements, we present three ideas of the assistant devices. The first one is a belt-type device which presents directions of the POV using vibration motors. The second one is a LED-type cylindrical device which presents directions of the POV using LED light sources. The third one is a robot-type device which presents directions of the POV using movement of the robot's head.

The belt-type device has an advantage that the broadcaster does not need to look the assistant device and can pay attention with his/her surroundings. It would highly satisfy requirement 1 and 2. We use a single line belt and the belt is bound around broadcaster's waist so that it can keep fashionability. Although the belt-type device can present the distinctive viewers' POV in the horizontal direction, it is difficult to present the distinctive viewers' POV in the vertical direction.

The LED-type cylindrical device has an advantage that it can present the distinctive POV both in the horizontal and vertical direction. However, the degree of achievement of the requirement 1 and 2 would be lower than the belt-type device because the LED-type device makes the broadcaster look the assistant device in order to check the distinctive viewers' POV.

The robot-type device specializes in requirement 3. It has an advantage that the robot can be a companion as if the broadcaster walked with a friend together and it would improve the broadcaster's experience. On the other hand, it has a disadvantage in terms of requirement 1 and 2 because it makes the broadcaster look and carry the device.



### 3 IMPLEMENTATION

We implemented prototypes of the assistant devices evaluate their effectiveness in an experiment. We used a 360-degree Internet live broadcasting system which was developed in our previous work [4] and added some functions to the broadcasting system so that it can handle the assistant devices. In this section, we describe the system architecture and prototypes of the assistant devices.

#### 3.1 Prototypes of the assistant devices

Described below are the new functions in this paper. The POV server analyzes the collected POV with the algorithm and send the distinctive viewers' POV to the assistant device. The assistant device has functions of presentation of the distinctive viewers' POV and reading comments in voice. The prototypes of the assistant devices are controlled by software on Raspberry Pi.

Figure 2 shows the prototypes of the assistant devices. The belt-type device has several vibration motors inside the belt at regular intervals. The broadcaster binds a belt around his/her waist and grasp the distinctive viewers' POV in a horizontal direction by the vibration. The LED-type device has several LED tapes which can light individually. The LED tapes are wound around a cylindrical can. The bottom of the can has a clip to fix it on something such as a laptop PC. The broadcaster can grasp the distinctive viewers' POV both in horizontal and vertical direction by the light. The robot-type device has a robot head and it can be turned up, down, right or left using two servo motors. The bottom of the robot head also has a clip to fix it on something. The broadcaster can grasp the distinctive viewers' POV both in horizontal and vertical horizontal direction by the direction of the robot head.

The reading comments in voice is a function which enables the broadcaster to confirm the viewers' comments in voice without looking a display device such as a smartphone and a laptop PC. In the previous work, the viewers' comments were displayed on a laptop PC. Since the prototypes of the assistant devices do not need the laptop PC, the broadcaster does not need to look the laptop PC. This function makes the broadcaster free from looking the laptop PC. In the outdoor environment, the broadcaster must be able to hear the surrounding sound for safety. We use a bone conduction headphone for hear the viewers' comments.

### 4 EVALUATION

#### 4.1 Experimental Procedure

We conducted two sets of experiments to evaluate the prototypes of the assistant devices. In each set of the experiments, there were one broadcaster and five viewers and the total experimental participants were 12 people. The location of the broadcasting was Takamatsu Pond in Morioka City, Iwate Prefecture, which was famous as a place where swans flied. Each broadcaster performed broadcasting for approximately 15 minutes and it was performed three times changing the assistant devices. The

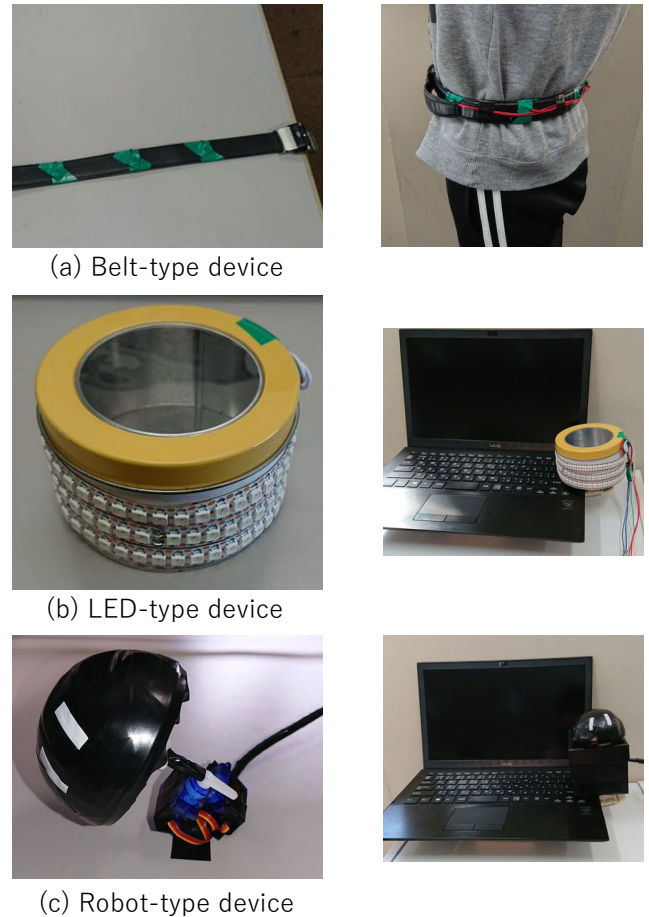


Figure 2: Prototypes of assistant devices

broadcaster walked around the pond and communicated with the viewers talking about the situation of the circumference. The viewers sent comments about the situation of the circumference of the broadcaster. We asked the broadcaster to stop at a safe place when checking the assistant device for his/her safety and to reply the comments from the viewers as many times as possible.

After three times of the experiments changing the assistant devices, we interviewed to the broadcaster about the presentation of the distinctive viewers' POV. In the interview, we asked about "Awareness of the presentation of the distinctive viewers' POV", "Awareness of the POV direction" and "Usefulness of the POV". Moreover, we asked the broadcaster to evaluate about "understandability of the POV direction", "Easiness of the POV check", "Utilization degree of the POV" and "Do you want to use it again?" on a scale of one to five. After that, we asked the broadcaster to give his/her impression about the assistant device through the whole experiment.

#### 4.2 Experimental Results

Table 1 shows the experimental result of the interview after the experiment. The presentation of the distinctive viewers' POV by the belt-type device was 24 times in 1st experiment and 23 times in 2nd experiment. The presentation by the LED-type device was 18 times in 1st experiment and 21 times in 2nd experiment. The

Table 1: Result of the interview after the experiment

Presentation Method	Experiment	Presentation Times	Awareness of the POV	Awareness of the POV direction	Usefulness of the POV
Belt-type	1st	24	24 (100%)	19 (79.2%)	14 (58.3%)
	2nd	23	23 (100%)	20 (87.0%)	11 (47.8%)
LED-type	1st	18	17 (94.4%)	8 (44.4%)	9 (50%)
	2nd	21	19 (90.5%)	12 (57.1%)	10 (47.6%)
Robot-type	1st	21	21 (100%)	18 (85.7%)	16 (76.2%)
	2nd	28	28 (100%)	24 (85.7%)	18 (64.3%)

Table 2: Result of the comparison

Presentation Method	Understandability of the POV direction	Easiness of the POV check	Utilization degree of the POV	Do you want to use it again?
Belt-type	3.5	5	3.5	4
LED-type	2	2.5	3	2
Robot-type	4	4	3	4

presentation by the robot-type device was 21 times in 1st experiment and 28 times in 2nd experiment. There was no significant difference among the assistant devices and it could be fair comparison. In terms of the awareness of the POV, they were aware of almost all of the POV presentations. In terms of the awareness of the POV direction, the POV presentations by the belt-type and robot-type devices could be understandable about 80%. However, that of the LED-type device was remarkably low, that is, around 50%.

Table 2 shows the experimental result of the comparison of the assistant devices. The belt-type and robot-type devices achieved scores above average in all comparison items. This result showed these devices were useful to the broadcaster. However, the LED-type device was low score in most comparison items without "utilization degree of the POV". The LED-type device especially would have problems of the usability.

We got the comments from the broadcaster written by a free format about the assistant devices. In terms of positive comments about the belt-type device, "it reduced labor and time for checking the POV" and "it can be equipped inconspicuously". In terms of negative comments about the belt-type device, "it was difficult to understand what the viewers were looking at" and "the direction of the POV was a little vague". These results showed the belt-type device was usable but there was an issue with understandability of the POV direction. The LED-type device got only negative comments. They were "difficult to check which LED were glowing", "it was difficult to understand which LED were growing" and "it took time and effort to check the glowing LED". These results showed the LED-type device was not suitable outdoors in the daytime. In terms of positive comments about the robot-type device, "it was easy to understand the direction of the POV", "I felt friendship for the robot", "it was cute" and "The direction of the POV was

best understandable". In terms of negative comments about the robot-type device, "I was worried about what surrounding people thought of my equipment" and "it took time to wait for the robot movement". These results showed the robot-type device gave friendship feeling to the broadcaster and it had an advantage in understandability of the POV direction though it stood out outside and the delay time was existed to present the POV.

## 5 CONCLUSION

In this paper, we proposed assistant devices for presentation of the distinctive viewers' POV in 360-degree Internet live broadcasting. we study and developed three prototypes of the assistant devices; a belt-type device which presented directions of the POV using vibration motors, a LED-type device which presented directions of the POV using LED light sources cylindrically and a robot-type device which presented directions of the POV using movement of the robot's head. We conducted an evaluation experiment. From the experiment, we found the broadcasters did not prefer assistant devices in visual form and they had a positive impression of the robot-type device.

For the future work, we will improve the robot-type device to make it more understandable and friendlier. We also study other methods to present the distinctive viewers' POV to the broadcaster using mixed reality (MR) technologies without disturbing the broadcasting.

## ACKNOWLEDGEMENT

This work was supported by JSPS KAKENHI Grant Number JP20K11794.

## REFERENCES

- [1] Roel Vertegaal: The GAZE groupware system: mediating joint attention in multiparty communication and collaboration, CHI '99 Proceedings of the SIGCHI conference on Human Factors in Computing Systems, pp.294-301 (1999).
- [2] David M. Grayson, Andrew F. Monk: Are you looking at me? Eye contact and desktop video conferencing, ACM Transactions on Computer-Human Interaction (TOCHI) Volume 10 Issue 3, September 2003, pp.221-243 (2003).
- [3] Masaya Takada, Dai Nishioka, and Yoshia Saito: A Detection Method of Viewers' Interests Based on POV for 360-Degree Internet Live Broadcasting in Mobile Environment, IEEE GCCE OS-VDP: 2D/3D Video Data Distribution and Processing (2019).
- [4] Masaya Takada and Yoshia Saito: A Study on Presentation of Viewers' Interests based on POV Analysis in Mobile 360-degree Internet Live Broadcasting, International Workshop on Informatics (IWIN2020), pp.3-8 (2020).

# An Efficient Large-Scale Video-on-Demand System on Edge Computing Environments

Satoru Matsumoto\* and Tomoki Yoshihisa\*

\*Cybermedia Center, Osaka University, Japan  
{smatsumoto, yoshihisa}@cmc.osaka-u.ac.jp

**Abstract**—Due to the recent increase in the users of video-on-demand (VoD) services, many clients such as smart phones or laptop computers request video data to a video distribution server. Such large-scale VoD systems utilize the edge computing technology to distribute the communication load and the processing load on the video distribution server. In most of the VoD systems utilizing edge computing, the edge servers receive a part of the video data from the video distribution server and caches them for other transmissions. However, the edge servers can receive them before receiving the requests from the clients (pre-cache). Moreover, the edge servers can transmit pre-cached video pieces to other edge servers. In this paper, we propose an efficient large-scale VoD system on edge computing environments. Our evaluation results revealed that our proposed system reduces the probability that the transmissions overlap with other transmissions and increases the shortest arrival interval of the clients that the system can work by 26% compared with the conventional system in the simulated situation.

**Keywords:** Streaming media distribution, interruption time, webcasting, Internet broadcasting, cloud computing

## 1 INTRODUCTION

Recently, video-on-demand (VoD) services such as YouTube or Netflix are widely used. In most VoD services, the clients request the video data to the video distribution server. The video distribution server sends the video data pieces sequentially to the clients so that they can play the video while receiving the pieces. When the clients cannot finish receiving each piece before starting playing it, the video playback is interrupted. A longer interruption time more annoys the viewers.

In large-scale VoD systems, many clients request the video data and thus the transmissions of data pieces frequently overlap with other transmissions. If the times required for transmissions increase and the transmissions continue to overlap with others, the interruption times also continue to lengthen. Therefore, existing schemes for interruption time reduction aim to reduce the transmission time to avoid the overlapping of transmissions. Major techniques for this are pre-caching ([1]–[4]), redistributions of data pieces ([5]–[7]), etc. Unfortunately, these traditional approaches cause the communication and processing loads on the clients. Such extra loads decline the users' operability of the clients and further consume the batteries if they are mobile devices.

Edge computing is one of the approaches that relief the computational loads on the clients since the edge servers, i.e.,

the servers on the edge of the network and geometrically close to the clients, are often managed by CDN companies such as Akamai or Cloudflare. In most of the VoD systems utilizing edge computing, the edge servers receive some pieces from the video distribution server and caches them for other transmissions. However, the number of the pieces that the edge servers need to send decreases by adopting the above both techniques (pre-caching and redistributions) to the edge servers.

In this paper, we propose an efficient large-scale VoD system on edge computing environments. Our proposed system adopts a distributed edge caching scheme. In our proposed system, the edge servers store some pieces before starting the VoD service. When a new client requests a video data, the video distribution server selects an edge server for sending the pieces of the requested video data to the client. The edge server sends its stored pieces to the client. After finishing sending the stored pieces, the video distribution server sends the subsequent pieces to the client. The contributions of the paper are (1) the increase of the number of the clients that the system can accommodate, (2) the proposition of a large-scale video-on-demand system, (3) the confirmation of the effectiveness of the proposed system.

The paper is organized as follows. Some related work are introduced in Section 2. The proposed scheme is explained in Section 3, and evaluated in Section 4. Finally, we conclude the paper in Section 5.

## 2 RELATED WORK

Many studies focus on fast data reception for VoD services.

### 2.1 Pre-caching Techniques for VoD Services

Abuhadra et al. proposed a proactive caching technique for mobile devices [1]. The probability that the clients encounter interruptions decreases by sending more video data to the mobile devices while their network connections are available because they sometimes disconnect from the network. The proposed technique reduces in-network transmission delays by caching the video data. Feng et al. found the optimal cache placement for the system with wireless multicasting [2]. My research group proposed a broadcasting method for pre-caching video data pieces predicting the video data that the client will play [3]. Coutinho et al. proposed a proactive caching technique for DASH video streaming [4]. In the proposed technique, the clients select a proxy caching server based on the network conditions. However, these pre-caching techniques for VoD services require the clients' storage capacity.

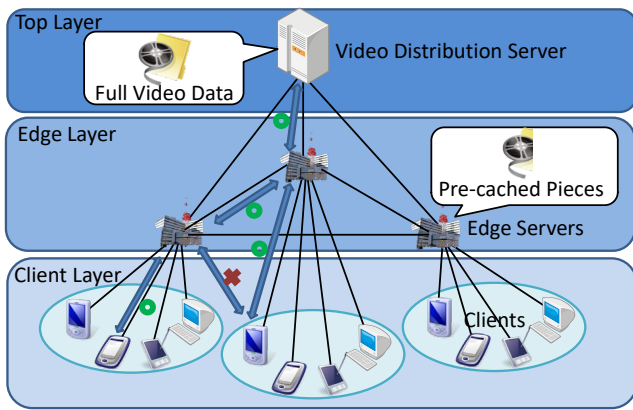


Figure 1. Assumed Environment

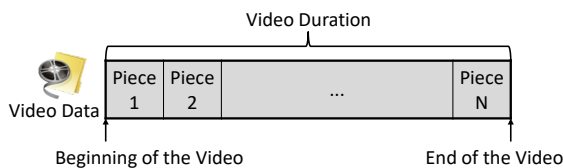


Figure 2. Video Data Division

## 2.2 Redistribution (Peer-to-Peer Sharing) Techniques for Video-on-Demand Services

Sheshjavani et al. proposed a peer-to-peer data sharing mechanism for VoD services [5]. In the proposed mechanism, the clients manage their buffer map. The buffer maps inscribe the received video data pieces and non-received pieces of each client. In the mechanism, the clients exchange the pieces based on the buffer map to receive the pieces that each client does not have. In the method proposed by Zhang et al., the clients send the pieces considering the bandwidth consumption [6]. Fratini et al. analyzed the efficiency of using replicated video servers [7].

However, similar to the pre-caching techniques, these redistribution techniques cause communication and processing loads on the clients. Such extra loads decline the users' operability of the clients and consume the batteries if they are mobile devices.

## 3 PROPOSED SYSTEM

In this section, we explain our proposed system.

### 3.1 Proposed System Architecture

Figure 1 shows the assumed system. The system consists of three layers, the top layer, the edge layer, and the client layer. One video distribution server is in the top layer. CDN (Contents Delivery Network) companies or VOD service companies provide the machines in the edge layer. The edge layer includes some edge servers. The clients are in the client layer. Similar to other researches for the VoD systems utilizing edge computing, the networks for these three layers are application layer networks and we assume that the influences of the underlaying session/transport layers are sufficiently small.

The single video distribution server has full video data for all the videos and connects to the Internet. The edge servers also connect to the Internet and can communicate with the video distribution server. They can store a part of some video data (pieces). The clients connect to the geometrically closest edge server and can communicate with the edge server. The clients request the video data to the video distribution server. Each video data are divided into some pieces, as shown in Figure 2. The clients receive the requested video data pieces from the video distribution server and the edge servers.

The assumed system model is general and practical. One of the applications is that the video distribution server provides the videos to the clients in a prefecture, and each edge server serves for each region in the prefecture. For example, there are eight local regions in Osaka, Japan. In this case, the number of the edge servers is eight, and the edge servers provides 100 videos.

### 3.2 Target Issue

We explain our target issue in this subsection.

#### 3.2.1. Issues in Conventional Systems

In the VoD systems, the video playback is interrupted when the clients cannot finish receiving each piece before starting playing it. A longer interruption time more annoys the viewers. Here, the interruption time means the total time that the playback is interrupted while a client is playing the video. In the VoD system in that a video distribution server distributes the video data to the clients, the communication load and the processing load for the distribution concentrate on the server. Therefore, in the cases that such VoD system hold many clients (large-scale) and the clients frequently requests to play the videos to the video distribution server, the server is easy to overload. The overloading results in long video data transmission times and causes long interruption times. In most of the VoD systems utilizing edge computing, the edge servers receive pieces from the video distribution server and caches them for other transmissions.

#### 3.2.2. Pre-Caching Pieces

The edge servers can receive pieces from the video distribution server before the requests for them come from the clients. That is, the edge servers can pre-cache the pieces. The pre-caching can reduce the communication load and the related processing load on the video distribution server and the edge servers. This is because they receive the pieces without the requests for the pieces and can receive them before starting the VoD service. Pre-caching more pieces reduce more loads, but require more storage capacity on the edge servers.

#### 3.2.3. Redistributing Pieces

Moreover, the edge servers can redistribute pieces to other edge servers. The redistribution can distribute the communication load and the related processing load on the video distribution server to the edge servers because the edge servers send the pieces instead of the server. However, if an



edge server frequently redistributes the pieces, the loads concentrate on the edge server, results in long interruption time. Therefore, the edge servers need to redistribute the pieces without causing the overloads on it.

### 3.2.4. Our Objective

Based on the discussion in this subsection, we aim to reduce the interruption time by adopting pre-caching and redistributing pieces on edge computing environments. For this, we propose a scheme which determines which pieces the edge servers should pre-cache and how to redistribute the pieces among the edge servers.

## 3.3 Proposed Scheme

Our proposed scheme runs on our proposed system architecture explained in Subsection 3.1. In the scheme, the edge servers pre-cache several preceding pieces of popular videos. When a client requests playing a video to the connected edge server, the edge server checks whether it has pre-cached the requested video already. If the piece that the client is going to receive is pre-cached, the client receives the piece from the connected edge server. If the piece is pre-cached by another edge server, the connected edge server receives it from the edge server and after that sends it to the client. If the pieces is not pre-cached by any edge server, the connected edge server receives it from the video distribution server and after that sends it to the client.

### 3.3.1. How to Pre-Cache Pieces

To solve the first issue, the edge servers pre-cache preceding several pieces of popular videos, i.e., the pieces that are close to the beginning of the videos. This is because the time to start playing the preceding pieces is early, and the possibility that the clients encounter interruptions is high. The preceding pieces of the videos that are frequently requested by the connected clients are pre-cached. To avoid redundant pre-caching, the edge servers do not pre-cache the pieces that are pre-cached by other edge servers. The edge servers do not cache the pieces that are transmitted to the clients except for the pre-cached pieces to reduce the required storage capacity for caching. The number of the pieces to be pre-cached is a parameter for the scheme.

For example, imagine the case that the VoD system provides 100 videos and owns eight edge servers. When the number of the pieces to be pre-cached is set by 10 per each video, each edge server pre-caches the preceding 10 pieces of the  $100/8=25$  videos.

### 3.3.2. How to Redistribute Pieces

To solve the second issue, in the cases that the edge server does not pre-cache the requested video, it requests the redistribution of the preceding pieces of the video to the edge server that pre-caches the requested video. This is because the edge layer and the client layer are separated in our proposed system, and thus, the clients cannot communicate with the edge servers that they do not belong to. Therefore, the edge

servers redistribute the pieces to other edge servers, not directly to the clients.

For example, imagine the case that the client 1 directly connects to the edge server 1 and requests the video 2. The preceding pieces of video 2 is pre-cached by the edge server 2. In this case, the edge server 2 redistribute the pre-cached pieces of the video 2 to the edge server 1 and the edge server 1 sends the received pieces to the client 1.

## 3.4 Flow of Procedures

In this subsection, we explain the flow of the procedures for the video distribution server, the edge servers, and the clients.

### 3.4.1. Video Distribution Server

The video distribution server has all the pieces of all the video data. When it receives a request for a piece, it sends the requested piece to the requesting edge server.

Moreover, the video distribution server manages the statistics of the VoD system and measure the popularity of the video data to determine which video data each edge server should pre-cache. This can be performed by calculating popularity of the video data. The video distribution server can calculate this by getting the information about the video requests from all the edge servers.

### 3.4.2. Edge Servers

When an edge server receives the request for the video from a client, it checks whether it pre-caches the first piece of the video or not. If the edge server pre-caches the piece, it sends the piece to the client. Otherwise, it finds the edge server that pre-caches the piece. If there is the edge server that pre-caches the piece, it requests the piece reception to the edge server and waits for the reception. When the reception completes, it sends the received piece to the client. If no edge servers pre-cache the piece, it request the piece reception to the video distribution server and sends it to the client when the piece reception completes. The edge server that receives the request of the video continue to this procedure until it sends the last piece. When it completes sending all the pieces of the video to the client, the flow for the request finishes.

### 3.4.3. Clients

When a client receives the first piece, it starts playing the piece. After playing the piece, the client try to continuously play the next piece and checks whether it has stored the next piece or not. If the next piece has already stored in its storage, it starts playing the next piece. Otherwise, it waits for the reception of the next piece. In this case, interruption occurs. The client continue to this procedure until finishing playing all the pieces.

## 4 EXPERIMENTAL EVALUATION

To check the performances of the proposed scheme, we measured the interruption time using our developed simulator.

### 4.1 Evaluation Setting

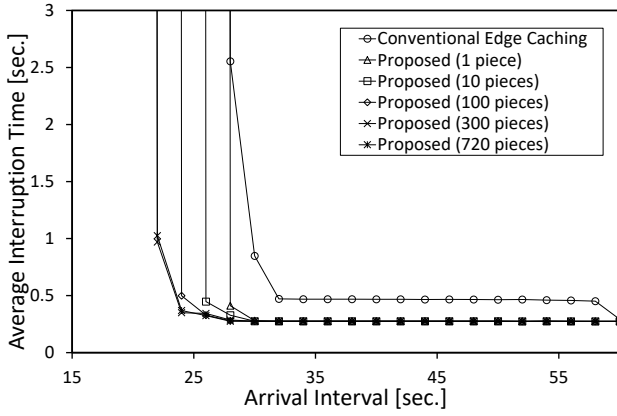


Figure 5. Average arrival interval and the interruption time

Based on the application example in Subsection 3.1, the number of the edge servers is eight, and the edge servers provide 100 videos. The bandwidth between each edge server and the clients is 100 [Mbps] considering a realistic situation. The bandwidth between the video distribution server and the edge servers is 600 [Mbps], and that among the edge servers is also 600 [Mbps], considering that these are in the backbone network. I set the same bandwidth to all the edge servers to make the experiments precisely understandable. The video distribution server can communicate with each edge server directly, and the edge servers can communicate with each other. The clients connect to the closest edge server.

The video duration is 60 [min.], and the bitrate is 5 [Mbps] based on the videos provided by practical services. The data amount of a piece is the same as the video data for 5 [sec.] based on HLS (HTTP Live Streaming [8]) and is 3125 [Kbytes]. The number of the pieces in each video data is 720.

We compare the proposed scheme with a conventional edge caching scheme, an often used caching technique for CDN. In the scheme, the pieces are cached at the edge servers, but not redistributed among them. The edge servers receive the pieces that need to be sent to the clients and are not cached from the video distribution server.

## 4.2 Influence of Arrival Interval

More frequent arrivals of the requests for playing the video data cause more transmission overlaps, and thus, the interruption time can continue to increase with a higher probability. Therefore, we investigate the influence of the arrival intervals of the clients' requests.

### 4.2.1. Interruption Time

Figure 5 shows the average interruption time under different arrival intervals. The horizontal axis is the arrival interval and the vertical axis is the average interruption time. In the legend, 'Conventional Edge Caching' indicates the average interruption time under the conventional edge caching scheme explained in the previous subsection. 'Proposed ( $J$  pieces)' indicated the average interruption time under our proposed scheme. In the scheme, each edge server pre-caches  $J$  pieces.

We can see that the average interruption times under each scheme are almost the same when the arrival interval is longer

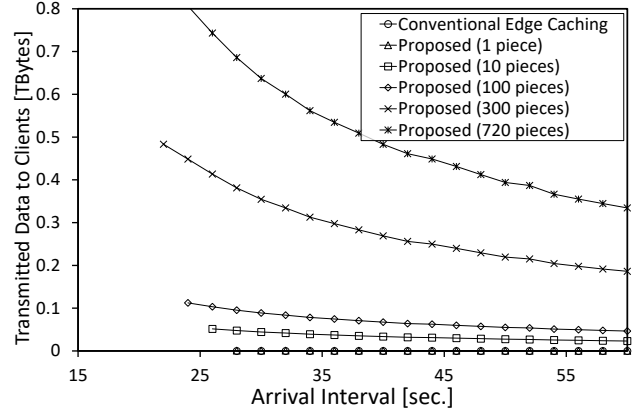


Figure 6. Average arrival interval and the transmitted data to the clients

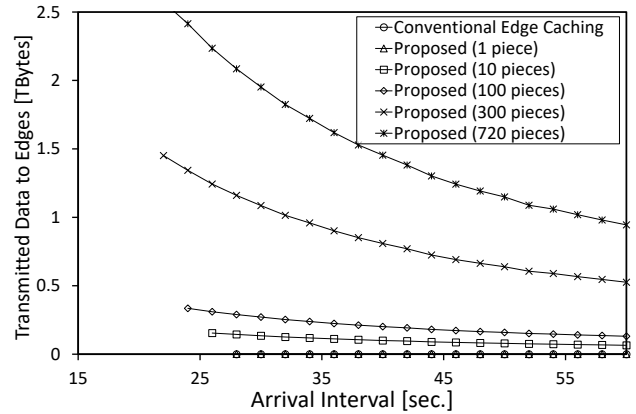


Figure 7. Average arrival interval and the transmitted data to the edge servers

than a certain value. This is because the transmissions does not overlap with others and the interruption time does not continue to increase. Under the conventional scheme, the average interruption time when the arrival interval is longer than 30 [s] is a little bit longer than that under our proposed scheme. This is because the bandwidth between the edge servers and the clients are not sufficient and the transmissions sometimes overlap with others. On the other hand, we can see that the average interruption times under all schemes suddenly increases when the arrival interval is shorter than a certain value. This is because the transmissions always overlap with the next transmission and the interruption time continue to lengthen. In such cases, the VoD system abandons and cannot provide their services.

For example, when the arrival interval is 25 [s], the conventional scheme cannot provide the service, but our proposed scheme can provide it and the average interruption time is approximately 325 [ms]. At the shortest, our proposed system can work even when the arrival interval is 22 [s]. The conventional method can provide service only when the arrival interval is longer than 30 [s] in this situation. Therefore, our proposed system can improve the shortest arrival interval under that the system can work 26% compared with the conventional system.

### 4.2.2. Edge Servers' Data Transmission

One of the indexes for the communication load and the processing load on the edge servers is the data amount

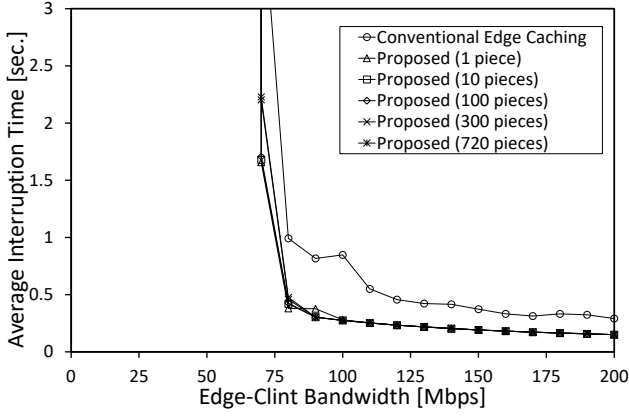


Figure 8. Edge-client bandwidth and the interruption time

transmitted to others. Therefore, we investigate the amount changing the arrival interval.

Figure 6 shows the total data amount transmitted to the clients by the edge servers. Since the average interruption time diverges when the arrival interval is excessively short, the lines stop at the shortest arrival interval under that the interruption time converges. The data amount decreases as the arrival interval increases because the number of the clients that receive data per time decreases. The data amount increases as the edge servers pre-cache more data because they need to transmit them instead of the video distribution server.

Figure 7 shows the total data amount transmitted to other edge servers. The result is similar to Fig. 6, the total data amount transmitted to the clients. But, the amount differs. The data amount transmitted to other edge servers is generally larger than that to the clients because the edge servers request the data transmissions to other edge servers in proportional to the number of the clients and each edge server respond to the requests from all the other edge servers.

We only show the total data amount transmitted to the clients because we confirmed that the total data amount transmitted to the edge servers is similar to this.

### 4.3 Influence of Edge-Client Bandwidth

A less communication bandwidth between the edge servers and the clients (edge-client bandwidth) cause more transmission overlaps, and thus, the interruption time can continue to increase with a higher probability. Therefore, we investigate the influence of the bandwidth between the edge servers and the clients.

Figure 8 shows the average interruption time under different client-edge bandwidth. The arrival interval is 30 [s]. The horizontal axis is the client-edge bandwidth, i.e., the bandwidth between each edge server and the clients. The vertical axis is the average interruption time. We can see that the average interruption times under our proposed scheme are almost the same when the edge-client bandwidth is larger than a certain value. Similar to the discussion for the previous subsection, this is because the transmissions does not overlap with others when the edge-client bandwidth is large. For the same reason as the previous subsection, the average interruption time under the conventional scheme is a little bit longer than that under our proposed scheme. Since the

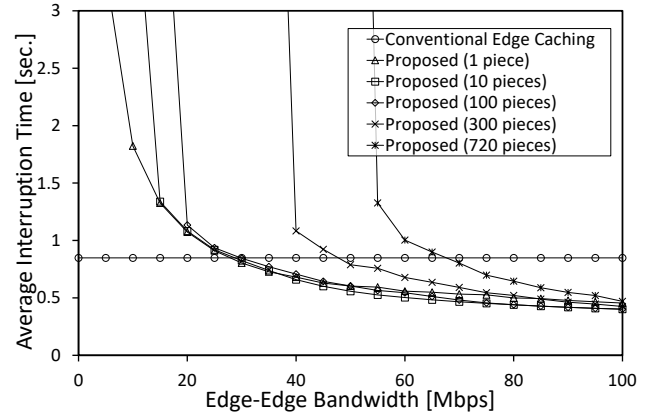


Figure 9. Edge-edge bandwidth and the interruption time

behavior between the edge servers and the clients are the same between our proposed scheme and the conventional scheme, the edge-client bandwidth under that the interruption time diverges is the same, approximately 75 [Mbps].

### 4.4 Influence of Edge-Edge Bandwidth

A less communication bandwidth among the edge servers (edge-edge bandwidth) cause more transmission overlaps. Thus, the interruption time can continue to increase with a higher probability.

Figure 9 shows the average interruption time. The arrival interval is 30 [s]. The horizontal axis is the edge-edge bandwidth. The average interruption time under the conventional scheme is constant even when the edge-edge bandwidth changes because the edge servers do not redistribute the pieces in the scheme. The average interruption time under our proposed scheme decreases as the edge-edge bandwidth increases because the edge servers can faster redistribute the pieces to another edge server. Moreover, the average interruption time continue to lengthen when the edge-edge bandwidth is smaller than a certain value because the transmissions overlap. The total data amount transmitted to the clients and the edge servers do not depend on the edge-edge bandwidth and these have a similar tendency as shown in Figs. 6 and 7.

### 4.5 Influence of Cloud-Edge Bandwidth

A less communication bandwidth between the video distribution server and the edge servers (cloud-edge bandwidth) cause more transmission overlaps. Therefore, we investigate the interruption time changing the cloud-edge bandwidth.

Figure 10 shows the average interruption time. The arrival interval is 30 [s]. The horizontal axis is the cloud-edge bandwidth. The average interruption time under the conventional scheme is longer than that under our proposed scheme because the bandwidth between the edge servers and the clients are not sufficient and the transmissions sometimes overlap with others as shown in Fig. 5. Similar to previous results, the average interruption time suddenly increases when the cloud-edge bandwidth is smaller than a certain value because the transmissions continue to overlap. The cloud-edge bandwidth under that the average interruption

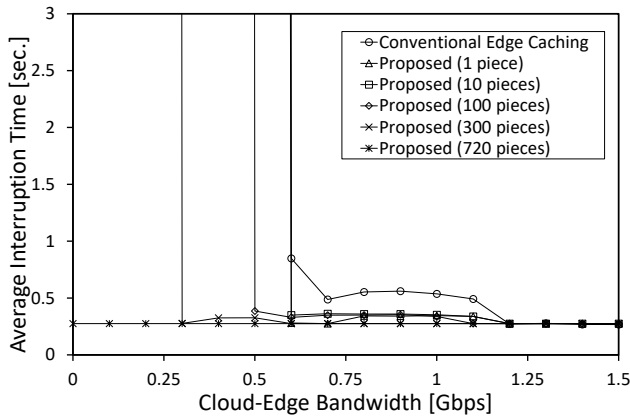


Figure 10. Cloud-edge bandwidth and the interruption time

time converges is larger as the number of the pre-cached pieces increases because the edge servers receive less pieces from the video distribution server as the pre-cached pieces increase. Since the edge servers pre-cache all the pieces when they pre-cache 720 pieces, the average interruption time does not change even when the cloud-edge bandwidth changes in this case.

## 5 CONCLUSION

In this paper, we proposed an efficient large-scale VoD system on edge computing environments. Our proposed system adopted the pre-caching and the redistribution techniques. Our proposed system reduces the probability that the transmissions overlap with other transmissions and increases the maximum number of the clients of that interruption time converges. Our simulation evaluation revealed that our proposed system can transmit the video data more than the conventional scheme. Our proposed system can improve the shortest arrival interval under that the system can work by 26% compared with the conventional system in the simulated situation.

In the future, we will consider the popularity of the videos and adopt the dynamic caching technique to the edge servers. Moreover, we will consider the communication loads on the edge servers and the distribution server.

## ACKNOWLEDGMENTS

This work was partially supported by JSPS KAKENHI Grant Numbers JP21H03429, JP18K11316, and by G-7 Scholarship Foundation.

## REFERENCES

- [1] R. Abuhadra and B. Hamdaoui, "Proactive In-Network Caching for Mobile On-Demand Video Streaming," in Proc. IEEE International Conference on Communications (ICC), pp. 1-6, 2018.
- [2] H. Feng, Z. Chen, H. Liu, and D. Wang, "Optimal Cache Placement for VoD Services with Wireless Multicast and Cooperative Caching," in Proc. IEEE Wireless Communications and Networking Conference (WCNC), pp. 1-6, 2018.
- [3] Y. Gotoh, T. Yoshihisa, and M. Kanazawa, Y. Takahashi, "A Broadcasting Protocol for Selective Contents Considering Available Bandwidth," IEEE Transactions on Broadcasting, vol. 55, no. 2, pp. 460–467, 2009.
- [4] R. Coutinho, F. Chiariotti, D. Zucchetto, and A. Zanella, "Just-in-time proactive caching for DASH video streaming," in Proc. Mediterranean Ad Hoc Networking Workshop (Med-Hoc-Net), pp. 1-6, 2018.
- [5] A.G. Sheshjavani, B. Akbari, and H.R. Ghaeini, "An Adaptive Buffer-Map Exchange Mechanism for Pull-based Peer-to-Peer Video-on-Demand Streaming Systems," Springer International Journal of Multimedia Tools and Applications, vol. 76, no. 5, pp. 7535–7561, 2016.
- [6] Y. Zhang, C. Gao, Y. Guo, K. Bian, X. Jin, Z. Yang, L. Song, J. Cheng, H. Tuo, and X.M. Li, "Proactive Video Push for Optimizing Bandwidth Consumption in Hybrid CDN-P2P VoD Systems," in Proc. IEEE International Conference on Computer Communications (INFOCOM), p. 2555-2563, 2018.
- [7] R. Fratini, M. Savi, G. Verticale, and M. Tornatore, "Using Replicated Video Servers for VoD Traffic Offloading in Integrated Metro/Access Network," in Proc. IEEE International Conference on Communications (ICC), pp. 3438-3443, 2014.
- [8] RFC 8216, HTTP Live Streaming, <https://tools.ietf.org/html/rfc8216>, 2017.

Session 5:  
Security and Algorithm  
( Chair: Tomoya Kitani )



# A Design of Plausibly Deniable Distributed File Systems

Ryouga Shibazaki<sup>†</sup>, Hiroshi Inamura<sup>‡</sup>, and Yoshitaka Nakamura<sup>\*</sup>

<sup>†</sup>Graduate School of Systems Information Science, Future University Hakodate, Japan

<sup>‡</sup>School of Systems Information Science, Future University Hakodate, Japan

<sup>\*</sup> Faculty of Engineering, Kyoto Tachibana University, Japan

{g2120017, inamura}@fun.ac.jp, nakamura-yos@tachibana-u.ac.jp

**Abstract** - Data protection has become an important issue in Internet services. In storage systems, conventional methods such as full disk encryption are generally used, but this alone cannot protect against forced attacks of key disclosure. PDE (Plausibly Deniable Encryption), which enables the denial of the existence of confidential information, has been proposed, and by disclosing the decoy key, it has become possible to protect the user from the force to disclose the key. It is an issue to be considered that the main memory is attacked at runtime due to the use in the cloud and the spread of virtualization technology. Therefore, we are proposing PTEE FS that realizes an encrypted file system using the concept of PDE in a trusted execution environment (TEE). To provide the resistance to exploit the knowledge from the use of disclosed decoy key, we introduce FID unification mechanisms. Regarding the performance of PTEE FS, we will evaluate the estimated performance given by the overhead of using TEE by using a model that imitates actual use on the cloud and using file synchronization between the server and client as the actual use model on the cloud.

**Keywords:** Plausibly Deniable Encryption, OS Security, Trusted Execution Environment.

## 1 Background

Leakage of confidential data related to privacy endangers the privacy of data owners and leads to the loss of social credibility of the leaked organization, so protection of such data has become an important issue. Traditional methods such as full disk encryption are commonly used in storage systems, but these methods make it difficult to maintain confidentiality when access to computer hardware or administrator privileges are stolen by an attacker. On the other hand, Plausibly Deniable Encryption (PDE), which is a new concept of encryption, has been proposed[1]. PDE protects confidential information sufficiently by allowing the existence of information to be denied. By disclosing the decoy key, PDE protects against the extortion of the decryption key by an attacker. While admitting that the encrypted file system exists in the system, the attacker is given the decoy key to access the decoy area, but the existence of the hidden area and its contents are kept secret. From the perspective of storage system configuration, PDE's existing research primarily protects sensitive information in persistent storage, and it is assumed that the main storage, which controls the existence of confidential information at runtime, will not be attacked. As an attack on the main memory, a memory inspection attack is assumed in this pa-

per. This is an attack that illegally takes a snapshot of the main memory and obtains confidential information.

So far, the purpose of this study is to construct an encrypted file system that is resistant to attacks not only on the permanent storage device but also on the main storage device and that can deny the existence using the concept of PDE. We proposed a system using Intel SGX as a hardware-protected execution environment in the realization of an encrypted file system[2].

In this paper, we examine countermeasures against attacks that are established on the premise that the attacker knows the existence of the decoy key and decoy data for PTEE FS (PDE with Trusted Execution Environment File System). As an evaluation of the processing time in normal access of PTEE FS, a model that imitates the actual use on the cloud is used, and the performance is evaluated in consideration of the overhead due to the use of TEE. In addition, as the processing time of the program started on demand, the performance is evaluated in consideration of the additional latency due to the FID merge processing described in Chapter 5. In the evaluation of processing time in normal access, file synchronization between server and client is used as an actual usage model on the cloud. In the evaluation of the processing time of the program started on demand, the local file created by referring to the existing research[3] by Leung et al. is used.

## 2 Related research and related technology

This section describes the concept of Plausibly Deniable Encryption, its application to file systems, and Intel SGX, which is being examined for application to the realization of attack resistance to main memory.

### 2.1 Plausibly Deniable Encryption

Plausibly Deniable Encryption (PDE) was proposed by Canetti et al. [1] as one of the encryption methods. Traditional disk encryption methods including full disk encryption, has the problem that it cannot be protected if the owner is forced to disclose the decryption key by an attacker. Therefore, PDE, which was proposed as one of the methods to protect the owner from the key disclosure extortion attack, enables the protection attack by using the decoy key. PDE is a characteristic of using a decoy key, which enables protection from key disclosure extortion attacks. As shown in Figure 1, PDE applies special encryption to confidential information that can be decrypted with both a decoy key and a private key, unlike conventional encryption. Decryption with the decoy key gives

the decoy plaintext, and decryption with the private key gives the original plaintext. When the legitimate user is attacked by an attacker forcing key disclosure, the user can give the decoy key to the attacker. Since the attacker thinks that the decoy key is the original private key, it allows the original confidential information unnoticed and kept secret.

On the other hand, the disadvantage is that the size of the ciphertext becomes extremely large, which may make the attacker suspicious of applying a special cipher. Furthermore, traces of confidential information may be obtained from the file system and the physical storage medium layer, etc., and considering these, it cannot be said to be a practical method. However, idea of PDE that decoy key gives decoy information and private key gives the confidential information can be used.

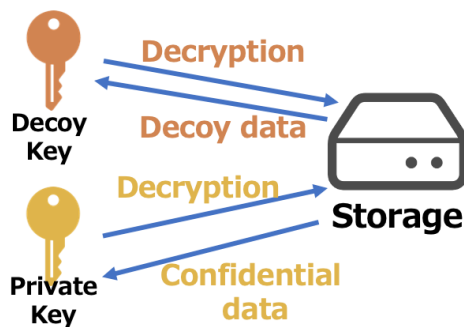


Figure 1: Overview of PDE

## 2.2 Applications of PDE concept

Using idea of PDE, a method was proposed to bring confidentiality using two types of techniques, steganography and hidden volume, instead of using simple encryption. First, a PDE method using the concept of steganography was proposed by Anderson et al.[4] and Chang et al.[5]. The basic idea is to hide confidential information in ordinary information. For example, confidential information is embedded and saved in a part of a large file such as an image file. In steganography, there is a risk that the confidential information will be overwritten when the file in which such confidential information is embedded is changed. In order to avoid overwriting confidential information, risk is alleviated by copying and saving multiple confidential information, but it has the disadvantage that the usage efficiency of the storage device deteriorates and a large amount of confidential information cannot be retained. PDE using hidden volume technology, has been proposed by Jia et al.[6] and Zuck et al.[7]. File system using hidden volume technology, creates a decoy volume on a storage device with a decoy key and a hidden volume with a private key. The decoy volume is placed throughout the storage device, and the hidden volume is usually placed from the hidden offset, which is the initial position of the hidden volume on the storage device, toward the end of the storage device. When using PDE file system using hidden volume technology, the user logs in in public mode or PDE mode and uses the file system. In public mode, user only operate decoy

volumes and in PDE mode, user can operate hidden volumes. When forced to disclose the key, the owner discloses the login password of public volume and the decoy key, so that protect hidden volume and the confidential data from the attacker. In the hidden volume technology, the existence of the hidden volume and the hidden offset are unknown in the system that operates the decoy volume, so the data stored in the decoy volume may overwrite the hidden volume.

## 2.3 Intel Software Guard Extensions

Intel Software Guard Extensions (Intel SGX)[8] is a CPU extension architecture provided by Intel Corporation. Intel SGX can perform processing that guarantees the confidentiality of data even if the privileged user or terminal administrator is not credible. As shown in Figure 2, Intel SGX create an encrypted area called Enclave on memory. Enclave provide a trusted execution environment (TEE) to enable program execution while maintaining data confidentiality provided at the hardware level. Intel SGX can protect the programs and data in the enclave from memory inspection attacks. Enclave are called using ECall from untrusted areas. Then, the result processed in the enclave is passed to the untrusted area using OCall. Enclave is executed by the CPU in a special mode which deny cannot be inspected and tampered by program outside Enclave. ECall and OCall can achieve confidentiality by deny access from cached address to Enclave's private memory by program outside Enclave. Intel Corporation provides the Intel Software Guard Extensions SDK as an environment for using Intel SGX technologies.

However, Enclave has a limit size that included both program and data, the size is about 100MB. Therefore, the content to be processed by the enclave must be minimized. For example, In the existing research[9] using Intel SGX by Ahemed et al., The policy is to keep only the private key and perform only the related processing in the enclave. A study measuring the performance of Intel SGX by Gjerdrum et al.[10] has shown that the overhead increases when the size of the buffer sent to the enclave exceeds 64 kB.

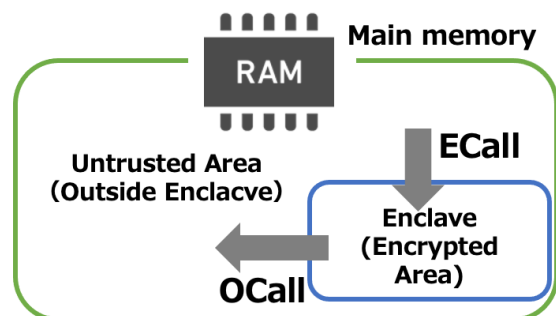


Figure 2: Overview of Intel SGX

## 2.4 Measured traffic for file server on the cloud

In the evaluation, we need to assume a usage model of file sharing on the cloud. Leung et al. [3] measured traffic for



two file-sharing servers used in NetApp data centers for three months. One of the servers was used by the marketing, sales and finance departments, and the other was used by the engineering department. This time, we referred to the statistical data of the servers used in each department of marketing, sales, and finance. This server received 364.3GB Read and 177.7GB Write access in 3 months. The ratio of Read, Write, and Delete requests was 540: 170: 1 in this order. The request size when accessing the file was about 70 % for less than 1kB, about 10 % for 1kB or more and less than 100kB, and about 20 % for those over 100kB.

### 3 Plausibly Deniable Distributed File Systems

The purpose of this research is to realize a plausibly deniable distributed file system that is resistant to key disclosure attacks and also resist memory inspection attacks in virtual environments.

#### 3.1 Base Design

In our research so far[2], we have designed a prototype of a distributed file system for key disclosure attacks as follows.

The basic idea of PDE is that using a decoy key or passphrase will give you information that is allowed to be disclosed, and using the original private key or passphrase will give you highly confidential information. In order to realize the basic idea of PDE, the proposed system provides a mechanism to switch the contents of the file handled based on the key and passphrase used for logging in to the file system.

PTEE FS server operates only the encrypted file, and does not operate the plaintext file, but PTEE FS client encrypts and decrypts the data and operate plaintext files. The server manages the decoy space and the hidden space. In the hidden area, highly confidential data such as access keys and passphrase for other systems that should not be leaked are stored. The decoy area does not include the data to be saved in the hidden area, and the data with low risk even if leakage occurs to the outside is saved. PTEE FS sever has the authorization control unit that determines whether the key sent from the client is decoy or authentic and switches the operation protects it using TEE (Trusted Execution Environment) and performs processing. The legitimate client PC and TEE are reliable areas, and the keys and passphrases used for user authorization are handled only in those areas. We use the NFS (Network File System) protocol with necessary modifications.

In propose configuration, it is necessary to switch the access destination into the decoy area and the hidden area by the key presented by the client and switch the structure of the file system. Code of the structure operation execute in TEE to prevent leakage and inspection by a snapshot of the main memory .

Since Intel SGX is used as the TEE, the confidentiality of the code for these structural operations can be maintained even when the attacker is a privileged user or terminal administrator. Therefore, this configuration can be resistant to infringement from snapshots of main memory when accessing the file system. However, with the TEE built using Intel SGX, there is a limit to the size of the enclave that can be

used, and there are some that cannot be used for kernel functions such as standard input / output in the enclave. In this research, we consider the security of the parts that are not protected by TEE, and propose the system configuration that protects them.

It is possible to obtain resistance to infringement from snapshots of persistent storage devices by performing processing such as filling empty areas on the file system with random bits as by Jia et al.[6].

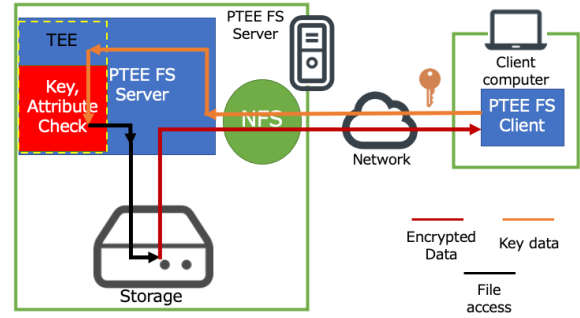


Figure 3: Data flow using TEE in the proposed system

## 4 Problem

We give a design resist attacks used knowledge from disclosure of decoy key, and obtain a practical prospect from the performance estimation when applied to cloud services. In addition to the attack methods we have examined so far, we describe attacks that use knowledge from the disclosure of decoy keys that have not been examined so far. Next, we explain the design of this system, and estimate the performance given by TEE when operating with the access pattern of the file synchronization service that is often seen in cloud storage services. In addition, we evaluate the performance of the system proposed in Chapter 5 when it is used in a typical workload when using cloud storage based on the existing research by Leung et al.[3].

### 4.1 Exploiting knowledge from the use of disclosed decoy key

We explain an attack that uses knowledge from the disclosure of the decoy key. When an attacker whose decoy key is disclosed can acquire the time series of attacker's access information to the decoy area by network traffic or a memory inspection attack on the server, the time series of access information to the hidden area by the private key by the legitimate user can be obtained, and the existence of the hidden area is revealed by comparing and collating these.

Regarding attacks using the knowledge of decoy key disclosure in PTEE FS, we will consider how the attacks are possible by monitoring the data exchange at the interface of TEE, and how to protect them. Figure 4 shows the data flow in the TEE interface. There are two interfaces, one between the network and TEE and the other between the persistent storage device and TEE. The information that can be observed in

each interface is defined as follows.

**TS1:** (TimeSeries1) In the operation time series between the network and TEE, the exchange of the modified NFS protocol is observed.

**TS2:** (TimeSeries2) In the operation time series between the persistent storage device and TEE, operation sequences such as fetch and store to the persistent storage device are observed.

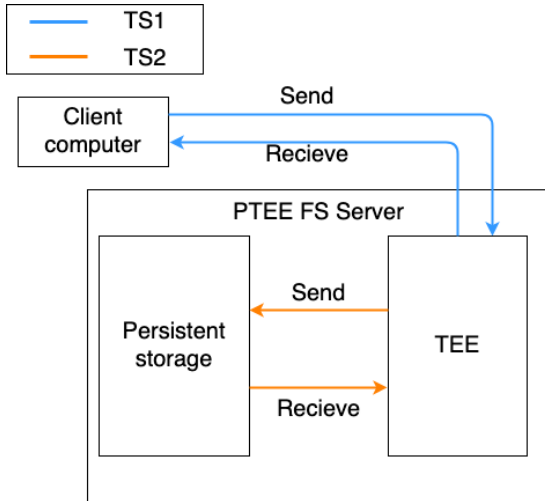


Figure 4: Data flow in TEE interface

Table 1 below shows examples of the contents observed by TS1 and TS2.

Table 1: Example of TS1 and TS2

TS1	Send	Recieve
	getattr File	(OK, Error) Result
	getattr Dir	(OK, Error) Result
	readdirplus Dir	(OK, Error) Result
	write File data	(OK, Error) Result
	read File	(OK, Error) Result data
TS2	Send	Recieve
	fetch ObjectID data	Result ObjectID (OK, Error)
	store ObjectID data	Result ObjectID (OK, Error)

TS1 is represented by the blue line in Fig. 1, and TS2 is represented by the orange line. TS1 and TS2 are arbitrarily generated by an attacker as TS1<sub>m</sub> and TS2<sub>m</sub> (m: malicious), and those generated by legitimate user operations are TS1<sub>l</sub> and TS2<sub>l</sub> (l:legitimate). At this time, the following two attacks can be considered from the information observable on the TEE interface.

**Attack Possibility 1:** Because of the attacker observing the difference between TS1<sub>m</sub> and TS1<sub>l</sub>, the existence of the hidden area is revealed

**Attack Possibility 2:** When TS2<sub>m</sub> is externally observable as an operation result of TS1<sub>m</sub>, it is possible to judge the match between TS1s from the unification of the pair of TS2<sub>m</sub> and TS2<sub>l</sub>, and the hidden area Existence is exposed

We place two assumptions are made as conditions for establishing "attack possibility 2".

**Attacker Assumption 1:** Correspondence between TS1 and TS2  $TS2 = TEE\_exposed\_func(TS1)$  can be estimated. This means that it is possible to associate the operation series from the NFS RPC time series to the operations for permanent storage device.

**Attacker Assumption 2:** It is possible to judge the match between the elements of TS2. In other words, it means that the operations on the persistent storage device can be identified and the unification can be observed.

Therefore, the following two are required to protect confidential information from attackers using the proposed method.

1. "Attack Possibility 1" is not established
2. Defend "Attack potential 2" by disabling either "Attacker Assumption 1" or "Attacker Assumption 2".

#### 4.1.1 Eliminating Attack Possibility 1

By encrypting the payload part of the RPC of the packet, which is a component of TS1, the difference other than the data size becomes unobservable, and the occurrence of "Attacker Possibility 1" can be prevented.

#### 4.1.2 Eliminating Attack Possibility 2

For "Attack Possibility 2", the following system configuration is adopted in order to prevent the "Attacker Assumption 1" from being established. As with the countermeasure for "Attack Possibility 1", the part related to RPC of the packet is encrypted. Regarding TS2, the data itself stored in the persistent storage device will be encrypted. In this configuration, the persistent storage device side assumes a general disk or a normal file system, so the object ID used when specifying the target in the persistent storage device is not protected from memory inspection attacks. The information obtained by the attacker at this time is the operation and object ID, the input / output timing to TEE, and the size of the encrypted part. It is necessary that the appearance pattern of the object ID in the IO traffic at the TEE does not provide any clue for attacker's tracking using TS2.

## 5 Design of PTEE FS

To solve the problems mentioned in Section 1, the object ID used by the attacker to specify in the persistent storage device observed by the TS2 should be the same between the decoy area and the hidden area as much as possible. To achieve this, there is a method of files that exist in any hidden area is embedded internally in one of the files in the decoy area.

Here, the hidden file is recognized as a free area by the system that handles only the decoy area.

Similarly, one directory in a hidden area should be embedded inside one of the directories in the decoy area. As a premise, where the decoy side or hidden side data exists in the persistent storage object actually read is recorded in the encrypted area of the persistent storage object. It is safely confirmed or operated in TEE which one should be accessed now.

## 5.1 FID unification procedure

The operation of embedding a hidden file inside a file in a decoy area in an appropriate directory structure is called FID unification processing. In the FID unification process, the same FID can be used by embedding the contents of the hidden area file in the file located in the appropriate directory structure of the decoy area. Embedding this file is called a merge operation. The FID unification process and merge operation are shown below. The FID unification process is used for initialization immediately after the proposed system is applied and for re-unification of unmerged files caused by file changes during operation. In order to merge the files in the hidden area into the files in the appropriate decoy area, the combination is searched to identify the appropriate location of the directory structure in the decoy area by the method shown in Algorithm1.

Algorithm1 operates as follows. First, get the path name list of all directories in the decoy area and the hidden area, and pass them to Function Search as an argument. In Function Search, the directory position of the decoy area, which is the starting point of the FID unification process, is determined from the combination of all the directories of the decoy area and the hidden area. The determination method is as follows. From the directory position of the decoy area that is the starting point, each directory of the decoy area and the hidden area has a one-to-one correspondence, and the following unification suitability evaluation is calculated by the operation shown in Algorithm2. The calculation method of the unification conformity assessment in Algorithm2 is explained in Section 5.1.1. The unification relevance evaluation consists of a mergeable flag and a conformance score. The mergeable flag is expressed by a boolean value indicating whether the directory combination can be merged, and when true, it indicates that the merge condition is satisfied. The calculation of the unification suitability evaluation is made into a memo, and when it is necessary to calculate the score of the same combination, it is called from the memo to shorten the calculation.

Among all combinations, the one with the maximum optimal score is selected from the ones for which the mergeable flag is true, and the directory position of the decoy area that is the starting point of the FID unification process is determined. If none of all combinations have the mergeable flag set to true, the one with the highest matching score is taken out and judged to be at risk based on that combination. Algorithm1 performs the processing up to this point and returns that the directory location of the decoy area that is the starting point of the FID unification process, or risk. The FID unification process recursively merges files or notifies the user of the

risk based on the result received from Algorithm1. There is a risk, that is, the mergeable flag obtained by Algorithm2 is not true because there are not enough decoy files in the decoy directory to be merged. Therefore, it calculates how many files should be added to which directory in the decoy area, and also notifies the user.

### 5.1.1 Conformance score

The conformance score integrates the conformance values for a specific file to be merged, and the larger the conformance score, the better the combination of the corresponding directories. A high match score means that the percentage of files in the decoy area where hidden area files are not embedded is high. In other words, if the conformance score is high, even if a new file on the hidden side is added or a file on the decoy side is deleted, there is a high possibility that the FID unification process can be performed only within the combination of the corresponding directories. It is used as a conformance score of the FID unification process. The average size of the files in the directory is the size of the decoy area as  $pSize$ , the size of the hidden area is as  $sSize$ . The number of files in the directory is the number of decoy areas as  $pNum$ , and the number of hidden areas as  $sNum$ . The calculation of the mergeable flag is  $(pSize/sSize + pNum/sNum)/2 \geq 2$ . The calculation of the conformance score is  $pNum/sNum$ .

## 6 Experiment

In this section, in order to consider the validity of the design of PTEE FS, the evaluation is performed using the verification case from the following two points.

### Processing time in normal access :

For the performance when applied to the cloud service of Section ??, first, we get the trace data of the file system acquired under assuming a realistic file group workload. The processing time is estimated applied our performance model [2] to the trace data got.

Regarding the FID unification process, we evaluate the effect of the smallest process among the FID unification processes that occurs when used in a typical workload. When the same usage as the file system using the existing PDE concept is used, it is the most called process in the proposed method, and the impact on the user is significant. So, we evaluate the effect of the additional latency to gave by FID unification process.

Create an environment that assumes the use case described in Section 6 of the proposed system, operate the FID unification process under that environment, and perform an evaluation experiment.

### 6.1 Experimental method

We prepared a decoy area directory and a hidden area directory according to the workload of the use case, and performed the FID unification processing. For the decoy area directory, referring to the existing research[3] by Leung et al., We prepared 70% for files with file sizes from 1 byte to 1 kB, 10%

**Algorithm 1** Algorithm to search for the best directory combination

---

```

1: function Serach(secretDirs, publicDirs)
2:   if secretDirs.length > publicDirs.length then ▷ If the hidden area has more directories than decoy area, no search
   is performed because there is no matching pattern.
3:     result  $\leftarrow$  noMatch
4:     return result
5:   allMatch  $\leftarrow$  allPermutaitionPatern(publicDirs) ▷ Calculate and substitute permutation patterns for directories in
   all decoy areas
6:   for i = 1,  $\dots$ , allMatch.length do ▷ Repeat the process for the number of allMatch
7:     for j = 1, secretFileNum do ▷ Repeat the process for the number of file in hidden area
8:       if resultMemo[j][allMatch[i][j]] = null then
9:         score  $\leftarrow$  CheckMatchDir(secretDirs[j], publicDirs[allMatch[i][j]]) ▷ Get the mergeable flag and
   optimal value for a combination of a directory in a decoy area and a directory in a hidden area
10:        resultMemo[j][allMatch[i][j]]  $\leftarrow$  score ▷ Save the score you have done once in a memo
11:      else
12:        score  $\leftarrow$  resultMemo[j][allMatch[i][j]] ▷ When the same combination appears, call it from the memo
13:        throughScore[i].optimal  $\leftarrow$  score.optimal ▷ Accumulate scores in the current permutation pattern
14:        if score.conform = false then
15:          throughScore[i].conform  $\leftarrow$  false
16:          throughScore[i].optimal  $\leftarrow$  -1
17:        if max(throughScore[i].optimal)! = 1 then ▷ Check if there is a mergeable combination
18:          result  $\leftarrow$  argmax(throughScore.optimal) ▷ Get the permutation pattern with the highest conformance score
19:        else
20:          result  $\leftarrow$  noMatch
21:        return result

```

---

for files with a file size of 1 kB to 100 kB, and 20% for files with a file size of 100 kB or more. We prepared three types of files, 30 and 50, contained in one decoy directory. For the hidden directory, referring to the key management of pgp, it was decided that the public key and private key pair of public key authentication, which is asymmetric authentication, is assigned to each directory. Two types of files, 10 and 20, are prepared in one hidden directory. If the number of files is 10, there are 5 public / private key pairs, and if the number of files is 20, there are 10 public / private key pairs. In the actual experiment, assuming that user use so that the number of files on the hidden area side is sufficiently small in PDE file system. We experiment 2 pairs of decoy area directory and hidden area directories. One pair is that the number of files in the decoy area directory is 30 and the number of hidden area directories is 10. The other is that number of files in the decoy area directory was 50 and the number of hidden area directories was 20. We excute the FID unification processing in these 2 pairs and the execution time was measured.

## 6.2 Experiment environment

A computer with RSYNC and NFS Version 3[11] installed was used as the server and client for the experiment. Wire Shark was used to trace the traffic. The network bandwidth in the experimental environment was 6.90 MB/s.

## 7 Evaluation

The average time required for FID unification is 133.8 ms with 30 files in the decoy area and 10 files in the hidden area,

and 134.6 ms with 50 files in the decoy area and 20 files in the hidden area.

## 8 Conclusion

We improved our design of Plausibly Deniable Distributed File Systems to obtain resistant to key disclosure attacks. Two experiments were conducted and evaluated in terms of performance in order to validate the design. In the experiments and evaluations, we discussed the processing time for normal access in use cases applied to cloud services. In the file synchronization use case using rsync, the increased ratio in response time by the use of TEE is estimated with measured figure. The result is 0.010%. increase for whole operation, which is considered to be acceptable overhead by TEE. To provide the resistance to exploit the knowledge from the use of disclosed decoy key, we added new functionalities of the FID unification as a countermeasure to memory inspection attacks. The processing time of the FID unification process invoked on demand was tested and evaluated using a program implemented in python.

The processing time of the FID unification process is 1133.8 ms in an environment with 30 files in the decoy area and 10 files in the hidden area, and 134.6 ms with 50 files in the decoy area and 20 files in the hidden area. Therefore, the additional latency due to the FID unification process may be tolerated. However, in this evaluation, the cost of encryption processing is not added to the processing time. Examination of a performance model that includes these is a future work.

**Algorithm 2** Algorithm for calculating the unification aptitude score

---

```

1: function CheckMatchDir(secretDir, publicDir)
2:   publicFiles  $\leftarrow$  getAllFiles(publicDir)            $\triangleright$  Get the file entry for the target decoy area directory
3:   secretFiles  $\leftarrow$  getAllFiles(secretDir)            $\triangleright$  Get the file entry for the target hidden area directory
4:   publicFileSizeMean  $\leftarrow$  publicFiles.sumSize/publicFiles.fileNum  $\triangleright$  Get the average file size of the decoy area
   directory
5:   secretFileSizeMean  $\leftarrow$  secretFiles.sumSize/secretFiles.fileNum  $\triangleright$  Get the average file size of the hidden area
   directory
6:   if publicFileSizeMean/secretFileSizeMean > 1 then            $\triangleright$  Check if the average file size meets the conditions
7:     sizeScore  $\leftarrow$  true
8:   else
9:     sizeScore  $\leftarrow$  false
10:  if publicFiles.fileNum > secretFiles.fileNum then            $\triangleright$  Check if number of files meets the conditions
11:    fileNumScore  $\leftarrow$  true
12:  else
13:    fileNumScore  $\leftarrow$  false
14:  if sizeScore & fileNumScore then
15:    conform  $\leftarrow$  true
16:  else
17:    conform  $\leftarrow$  false
18:  if conform then            $\triangleright$  Check if both the average file size and number of files meets the conditions
19:    optimal  $\leftarrow$  publicFiles.fileNum/secretFiles.fileNum  $\triangleright$  If the mergeable flag is true, the optimum value is
   calculated.
20:  else
21:    optimal  $\leftarrow$  0
22:  return (conform, optimal)            $\triangleright$  Returns mergeable flag, conformance score

```

---

**REFERENCES**

- [1] Rein Canetti, Cynthia Dwork, Moni Naor, and Rafail Ostrovsky. “Deniable Encryption”. In Burton S. Kaliski, editor, *Advances in Cryptology — CRYPTO ’97*, Lecture Notes in Computer Science, pp. 90–104. Springer Berlin Heidelberg, 1997.
- [2] Shibazaki Ryouga, Inamura Hiroshi, and Nakamura Yoshitaka. “Design of Encrypted File System Using the Concept of PDE”. *Proceedings of the 82th National Convention of IPSJ*, Vol. 82, No. 1, pp. 103–104, 2020.
- [3] Andrew W. Leung, Shankar Pasupathy, Garth Goodson, and Ethan L. Miller. “Measurement and Analysis of Large-Scale Network File System Workloads”. In *2008 {USENIX} Annual Technical Conference ({USENIX} {ATC} 08)*, 2008.
- [4] Ross Anderson, Roger Needham, and Adi Shamir. “The Steganographic File System”. In *Information Hiding*, pp. 73–82. Springer, Berlin, Heidelberg, April 1998.
- [5] B. Chang, F. Zhang, B. Chen, Y. Li, W. Zhu, Y. Tian, Z. Wang, and A. Ching. “MobiCeal: Towards Secure and Practical Plausibly Deniable Encryption on Mobile Devices”. In *2018 48th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*, pp. 454–465, June 2018.
- [6] Shijie Jia, Luning Xia, Bo Chen, and Peng Liu. “DEFTL: Implementing Plausibly Deniable Encryption in Flash Translation Layer”. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, CCS ’17*, pp. 2217–2229, New York, NY, USA, 2017. ACM.
- [7] Aviad Zuck, Udi Shriki, Donald E. Porter, and Dan Tsafir. “Preserving Hidden Data with an Ever-Changing Disk”. In *Proceedings of the 16th Workshop on Hot Topics in Operating Systems, HotOS ’17*, pp. 50–55, New York, NY, USA, 2017. ACM.
- [8] “Intel® Software Guard Extensions (Intel® SGX)”. <https://www.intel.com/content/www/us/en/architecture-and-technology/software-guard-extensions.html>. (accessed 2021-06-11).
- [9] Rufaida Ahmed, Zirak Zaheer, Richard Li, and Robert Ricci. “Harpocrates: Giving Out Your Secrets and Keeping Them Too”. In *2018 IEEE/ACM Symposium on Edge Computing (SEC)*, pp. 103–114, Seattle, WA, USA, October 2018. IEEE.
- [10] Anders T. Gjerdrum, Robert Pettersen, Håvard D. Johansen, and Dag Johansen. “Performance of Trusted Computing in Cloud Infrastructures with Intel SGX:”. In *Proceedings of the 7th International Conference on Cloud Computing and Services Science*, pp. 696–703, Porto, Portugal, 2017. SCITEPRESS - Science and Technology Publications.
- [11] B. Callaghan, B. Pawlowski, and P. Staubach. “NFS Version 3 Protocol Specification”. <https://www.ietf.org/rfc/rfc1813.txt>, June 1995. (accessed 2019-12-24).



# A Study on Division Impossibility in the Lightweight N-party Secure Function Evaluation for Cloud-Edge Computing Applications

Yutaro Taki<sup>†</sup>, Shigeru Fujita<sup>‡</sup>, and Norio Shiratori\*

<sup>†</sup>Graduate School of Information and Computer Science, Chiba Institute of Technology, Japan

<sup>‡</sup>Faculty of Information and Computer Science, Chiba Institute of Technology, Japan

\* Research and Development Initiative, Chuo University

**Abstract** - Lightweight three-party secret function calculations include variance, addition, subtraction, constant multiplication, and multiplication. Logical operation protocols have been defined. The same protocols are defined for a lightweight N-party secret function, which is a generalization of the lightweight three-party secure function. However, both these methods do not define constant division and division protocols. On the  $\mathbb{Z}/p\mathbb{Z}$  ring with the prime  $p$  as the law, an inverse is always obtained because of the nature of the congruence formula.

**Keywords:** Secure Computation, N-party Secret Sharing, N-party Secure Function Evaluation, Proof

## 1 Introduction

Data security is essential to ensure privacy, and it is indispensable in cryptography and information system design, development, and operation. Extensive computations are required to secure a single file. In addition, leaving data with others for a long time, even when it is encrypted, increases the risk of its unauthorized decryption [1].

Therefore, data deposition in a cloud system operated by a third party is always insecure. Secret sharing methods, which mix random numbers into data, divide it, and store it in a distributed manner, have been effectively used in untrusted cloud systems [2]. A secret calculation method executes data calculations in a secretly shared state [3]. By combining these two, the result can be obtained while ensuring data confidentiality. Secretly shared data cannot be analyzed independently. In addition, if the number of data is less than a given number, an analysis is impossible, even if it leaks in the same way [4].

Therefore, the storage of private data is recommended. Many public institutions handle privacy data. However, ensuring adequate security in data storage can be expensive. The secret sharing/secret calculation method is expected to reduce such expenses [5].

A lightweight N-party secure function evaluation [6] is a method that can execute addition, subtraction, multiplication, and logic circuit calculation within the framework of secret sharing/calculation when the number of collusions between subjects is suppressed to a certain number or lower. So far, we have considered the execution of division. In this study, we show that division cannot be performed within the framework of the proposed method and highlight the necessity for considering a new method.

A lightweight N-party secure function evaluation targets the remainder of a large prime number. This transformation

is equivalent to a projective operation. Since division requires two projections, such a number cannot be determined as a result of this operation. Lightweight three-party secret function calculations include variance, addition, subtraction, constant multiplication, and multiplication. Logical operation protocols have been defined [7]. We also successfully define the same computational protocol is defined for a lightweight N-party secret function computation, which is a generalized version of the lightweight three-party secure function.

However, both methods do not define constant division and division protocols. In this paper, we also successfully define on the  $\mathbb{Z}/p\mathbb{Z}$  ring with the prime  $p$  as the law, the inverse is always obtained because of the nature of the congruence formula.

## 2 Related Work

In this chapter, we describe the basics of the N-party secret function calculation used in this study. Specifically, this example shows how the protocol works and the type of environment in which it takes place.

A typical use-case for N-party secret function calculation is systems where both security and availability are required. This includes fields such as finance and medicine. For example, statistical processing is performed on clinical research data [8], [9]. Some of the SDK is available as open source [10], [11].

Figure 1 shows the protocol of a system using the N-party secret function calculation.

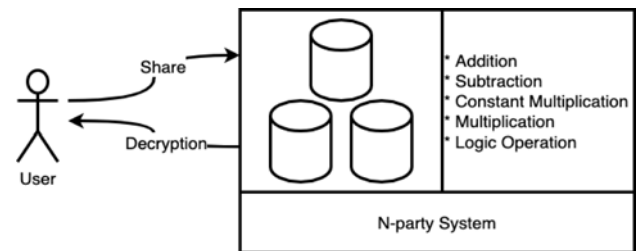


Figure 1: N-party System

The user and the system exchange values using the share and decryption protocols. In addition, the system can internally create a share of calculation results using addition, subtraction, constant multiplication, multiplication, and logic operation protocols.

Here, the concept of share is explained, including the basics of N-party concealment function calculations. First, the pa-

rameters  $n$  and  $k$  play important roles in N-party concealment function calculations. Specifically,  $n$  represents the number of entities with computing and communication capabilities, such as servers and electronic devices.

In addition,  $k$  defines the number of calculation entities that should be collected for the restoration. In other words, in the N-party concealment function calculation system, even if data leaks from  $k$  subjects, it does not lead to a leakage of the original data. In other words, a larger value of  $k$  increases safety, and a smaller value increases availability. Also, the relationship between  $n, k$  is  $n \geq k$  because  $k$  cannot exceed the number  $n$  of subjects. Furthermore, it is known that  $n \geq 2k - 1$  must be present for the multiplication protocol to work [6]. Further, subject  $i$  is expressed as  $P_i$ .

Shares are created using a share protocol and are a collection of multiple values. Specifically, the share of a certain value  $a$  sent to subject  $P_i$  is expressed as  $[a]_i$ . The values of this system can range from  $0 \dots p - 1$ , which are based on the prime number  $p$ . This is expressed as  $\mathbb{Z}/p\mathbb{Z}$  using the coset ring notation.

The share and the decryption protocols are shown below.

#### Share : k-out-of-n secret sharing

Input :  $a \in \mathbb{Z}/m\mathbb{Z}$   
Output :  $P_i([a]_i)$

1. Randomly select  $a_0, \dots, a_{n-2} \in \mathbb{Z}/m\mathbb{Z}$ .
2. Calculate  $a_{n-1} := a - \sum_{i=0}^{n-2} a_i$ .
3. Send to  $P_i$  as  $[a]_i := (a_i, \dots, a_{n-k+i})$  about  $i = 0, \dots, n - 1$ .

#### Dec : k-out-of-n secret decryption

Input :  $P_i([a]_i)$   
Output :  $a$  or  $\perp$

1.  $P_i$  discloses  
 $(\alpha_i, \dots, \alpha_{n-k+i}) := (a_i, \dots, a_{n-k+i})$ .
2. If there is at least one  $\alpha_i \neq a_i$  that is an  $a_i$  for  $\alpha_i$  corresponding to each subject  $P_i$  in  $i = 0, \dots, n - 1$ , it returns  $\perp$  to indicating an abnormality and terminates.
3. Calculate  $a = \sum_{i=0}^{n-1} a_i$ .

The share and decryption protocols allow users to exchange data with the N-party concealment function computing system. Further, as examples of the calculation protocols, the addition/subtraction and constant multiplication protocols are shown below.

#### Add/Sub : Create a $a \pm b$ share from an $a, b$ share

Input :  $P_i([a]_i, [b]_i)$   
Output :  $P_i([a \pm b]_i)$

- $P_i$  calculates  $[a \pm b]_i = (a_i \pm b_i, \dots, a_{n-k+i} \pm b_{n-k+i})$ .

#### CoMul: Create a $ca$ share

Input :  $P_i([a]_i), c$   
Output :  $P_i([ca]_i)$

- $P_i$  calculates  $[ca]_i = (ca_i, \dots, ca_{n-k+i})$ .

The multiplication protocol performs computation while communicating between entities. In order to show this, it is necessary to show the overall picture of the process, the process in each entity, and the allocation of the process. These contents cannot be written in abbreviated form. They are long and complex, and there is no space for them. The multiplication protocol has already been described in the previous study [6]. Each calculation in the N-party concealment function can be performed using protocols, such as addition/subtraction and constant multiplication, shown above.

However, the division protocol has not been defined. The next section covers why the division protocol is not defined, that is, why the division protocol is impossible.

### 3 Impossibility of division protocol

In this section, we discuss the impossibility of the division protocol. As mentioned in Chapter 2, share, decryption, addition, subtraction, multiplication of constants, multiplication, and logical operation protocols are defined for a lightweight three-party secure function evaluation and a lightweight N-party secure function evaluation. However, a division protocol is not defined among the four arithmetic operations.

In this section, we prove that there is no the impossibility of a division protocol because the lightweight three-party Secure Function evaluation and lightweight N-party secure function evaluations are performed in the modulo ring  $\mathbb{Z}/p\mathbb{Z}$  modulo ring based on the prime number  $p$ .

Specifically, we show that there is no one-to-one correspondence between the result of division on a modulo ring and the result of division on a set of integers.

#### 3.1 Preliminaries

For proving the absence of a division protocol, finite groups, modulo ring, Fermat's little theorem, and share and decryption protocols are explained in this section.

##### 3.1.1 Finite group and Residue Class Rings

In a finite group, set  $\mathbb{G}$  has only a finite number of elements. Lightweight three-party secure function and lightweight N-party secure function evaluations, are performed in residue-class rings, which are a type of finite group.

A modulo ring is a set of remainders obtained by dividing an arbitrary element  $a$  by a natural number  $n$ . For example, we can tell the time by collecting the readings between 0 and 23 o'clock and the data to be of 16 bits when values between 0 and to 65,535 are collected. Since the lightweight three-party and lightweight N-party secure function evaluations are divided by the prime number  $p$ ,  $p - 1$  is collected from 0. Thus, it can be consider the  $\mathbb{Z}/p\mathbb{Z}$  coset ring is a finite group because it has only  $p$  finite elements.



### 3.1.2 Fermat's little theorem

Let  $p$  be an arbitrary prime number. According to Fermat's little theorem, the following equation (1) holds for an integer that is not a multiple of  $p$ .

$$a^{p-1} \equiv 1 \pmod{p} \quad (1)$$

The case of  $p = 5$  is shown below.

$$1^4 = 1 \equiv 1 \pmod{5} \quad (2)$$

$$2^4 = 16 \equiv 1 \pmod{5} \quad (3)$$

$$3^4 = 81 \equiv 1 \pmod{5} \quad (4)$$

$$4^4 = 256 \equiv 1 \pmod{5} \quad (5)$$

In addition, for an integer  $a$  that is not a multiple of  $p$ , with  $p$  being a prime number, an  $x$  that satisfies the following equation (6) exists uniquely in the  $\pmod{p}$ .

$$ax \equiv 1 \pmod{p} \quad (6)$$

This  $x$  is generally called "the inverse element of  $a$  in  $\pmod{p}$ ." The proof that the inverse element's existence is proven below as *reductio ad absurdum*. Note that the constraints on  $x, y$  are  $1 \neq x, y \neq p-1$  and  $x \neq y$ .

*Proof.* Suppose we have  $x, y$  such as  $xa \equiv ya \pmod{p}$ . Then the equation  $a(x - y) \equiv 0 \pmod{p}$ . Since  $a$  is not a multiple of  $p$ , it becomes  $x - y \equiv 0 \pmod{p}$ .

$1 \neq x$  is a contradiction because  $y \neq p-1$  more  $x = y$  is not.  $\square$

Example of Division failure importantly, when the inverse is present on the  $\mathbb{Z}/p\mathbb{Z}$  coset ring, the quotient is always obtained using  $x \div y$ , where in  $1 \neq x, y \neq p-1, x \neq y$  of in  $x, y$ . As an example, find the quotient for  $p = 11, x = 3, y = 4$ .

$$4 \times 9 = 36 \equiv 3 \pmod{11} \quad (7)$$

$$3 \div 4 \equiv 9 \pmod{11} \quad (8)$$

This result is different from the integer division result.

$$3 \div 4 = 0 \quad (9)$$

### 3.1.3 Share and decryption protocols

Further, the share and decryption protocols of a lightweight N-party secure function evaluation are shown below. The important point here is that the input of the share protocol is the ring of remainders in  $\mathbb{Z}/m\mathbb{Z}$ , and the output of the decryption protocol is the ring of integers in  $a$ .

**Share :** k-out-of-n secret sharing

Input :  $a \in \mathbb{Z}/m\mathbb{Z}$

Output :  $P_i([a]_i)$

1. Randomly select  $a_0, \dots, a_{n-2} \in \mathbb{Z}/m\mathbb{Z}$ .
2. Calculate  $a_{n-1} := a - \sum_{i=0}^{n-2} a_i$ .
3. Send to  $P_i$  as  $[a]_i := (a_i, \dots, a_{n-k+i})$  for  $i = 0, \dots, n-1$ .

**Dec :** k-out-of-n secret decryption

Input :  $P_i([a]_i)$

Output :  $a$  or  $\perp$

1.  $P_i$  discloses  $(\alpha_i, \dots, \alpha_{n-k+i}) := (a_i, \dots, a_{n-k+i})$ .
2. If there is at least one  $\alpha_i \neq a_i$  that is an  $a_i$  for  $\alpha_i$  corresponding to each subject  $P_i$  in  $i = 0, \dots, n-1$ , it return  $\perp$  to indicate an abnormality and terminates.
3. Calculate  $a = \sum_{i=0}^{n-1} a_i$ .

These two protocols map a ring of integers onto a modulo ring and vice versa. Therefore, in the lightweight N-party secure function evaluation, there must be a one-to-one correspondence between the results of the ring of integers and the modulo ring.

## 3.2 proof

Finally, we prove the impossibility of a division protocol based on the above discussions.

*Proof.* Suppose we have the following division protocol:

**Div :** Create  $a \div b$  from the shares of  $a$  and  $b$

Input :  $P_i([a]_i, [b]_i)$

Output :  $P_i([a \div b]_i)$

1. some form of processing
2. some form of processing

Owing to the nature of the share protocol, the share  $P_i([a \div b]_i)$  of the division created by the decryption protocol is in the modulo ring.

In addition, because of the inverse nature of Fermat's little theorem, even if  $a \div b$  is a value that causes a remainder in the ring of integers, it is divisible. When  $P_i([a \div b]_i)$  created as a result, is input to the restoration protocol and returned to the ring of integers, its value will be different from the required  $a \div b$ .

Therefore, a division protocol is impossible to define.  $\square$

## 4 Conclusion

In the N-party secret function calculation, we were able to internally create a share of the calculation results using the addition, subtraction, constant multiplication, multiplication, and logic operation protocols. However, a division protocol was not defined. In this study, we showed that a division protocol is impossible in an N-party secret function computation based on the properties of finite groups, modulus algebras, Fermat's little theorem, and secret sharing and decryption protocols.

Specifically, in this paper, we showed that there is no one-to-one correspondence between the result of division on a modulo ring and the result of division on a set of integers.

In actual operations, the users must perform division themselves after the division processing is finally restored by the allocation decryption protocol. Our future work includes the study and implementation of a statistical process that includes commonly used divisions.

## REFERENCES

- [1] M Swathy Akshaya and G Padmavathi. Taxonomy of security attacks and risk assessment of cloud computing. In *Advances in big data and cloud computing*, pages 37–59. Springer, 2019.
- [2] Adi Shamir. How to share a secret. *Communications of the ACM*, 22(11):612–613, 1979.
- [3] Chuan Zhao, Shengnan Zhao, Minghao Zhao, Zhenxiang Chen, Chong-Zhi Gao, Hongwei Li, and Yu-an Tan. Secure multi-party computation: Theory, practice and applications. *Information Sciences*, 476:357–372, 2019.
- [4] Yfke Dulek, Alex B Grilo, Stacey Jeffery, Christian Majenz, and Christian Schaffner. Secure multi-party quantum computation with a dishonest majority. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 729–758. Springer, 2020.
- [5] Yongjune Kim, Ravi Kiran Raman, Young-Sik Kim, Lav R Varshney, and Naresh R Shanbhag. Efficient local secret sharing for distributed blockchain systems. *IEEE Communications Letters*, 23(2):282–285, 2018.
- [6] Taki Yutaro, Fujita Shigeru, Miyanishi Yohtaro, and Shiratori Norio. Lightweight n-party secure function evaluation with error detection. *Journal of Information Processing Society of Japan*, 59(10):1895–1902, oct 2018.
- [7] Chida Koji, Ikarashi Dai, Hamada Koki, and Takahashi Katsumi. A lightweight three-party secure function evaluation with error detection and its experimental result. *Journal of Information Processing Society of Japan*, 52(9):2674–2685, sep 2011.
- [8] Trial service of secure computation system san-shi<sup>®</sup> —toward safe and secure use of confidential data— —press release — ntt. <https://group.ntt/en/newsrelease/2018/08/08/180808a.html>. (Accessed on 07/20/2021).
- [9] Secure computation technology for analyzing data with the information concealed — nec. <https://www.nec.com/en/global/rd/technologies/201805/index.html>. (Accessed on 07/20/2021).
- [10] Dan Bogdanov, Sven Laur, and Jan Willemson. Sharemind: A framework for fast privacy-preserving computations. In *European Symposium on Research in Computer Security*, pages 192–206. Springer, 2008.
- [11] Privacy enhancing technology for data-driven business — sharemind. <https://sharemind.cyber.ee/>. (Accessed on 07/20/2021).

# Verification of Shell Script Behavior by Comparing Execution Log

Hitoshi Kiryu<sup>†</sup>, Shinpei Ogata<sup>†</sup>, and Kozo Okano<sup>†</sup>

<sup>†</sup>Faculty of Engineering, Shinshu University, Japan  
21w2025g@shinshu-u.ac.jp, {ogata, okano}@cs.shinshu-u.ac.jp

**Abstract** - Scripts written in Bash are widely used for task automation on UNIX OS such as Linux. These scripts are taken over after OS upgrades. Behaviors of the scripts can be changed by the upgrades. In order to deal with the behavior changes, developers have to inspect scripts in every release. However, it needs a lot of costs. This paper proposes a method to verify whether the script behavior is change between two different OS versions. It also detects the cause of the difference. An automated tool which is based on the method is also presented. In result of evaluation, we confirmed that the proposed method can verify the difference and detect commands that cause the difference in a simple example. In addition, it was found that it is difficult to detect the changes that does not appear in the standard output. In order to detect these commands as the cause, it is necessary to collect logs from a different area from the standard output.

**Keywords:** Behavior verification, Debug, Execution log, Shell script, Bash

## 1 INTRODUCTION

### 1.1 Background

Bash[1] is a typical shell language that runs on UNIX OSes such as Linux. These scripts written in shell are widely used for automating tasks. The behavior of the scripts written in the Bash may change due to OS updates and associated command upgrades. Therefore, when the OS is upgraded, developers must understand changes in the commands in the scripts and OS specifications. It costs a lot to debug scripts and each command in the scripts.

Various studies[2]–[5] have been conducted to localize bugs related to these problems. These studies proposed methods to identify statements that cause bugs by using bug reports, trace information, visualization. These studies aim to fix and identify bugs, while our purpose is to identify the causes of the differences in behavior.

The author’s group has been researching behavioral equivalence verification of programs[6]–[8]. In these studies, we verified the differences in the behavior of modified programs as shown on the left side of Figure 1. The problem that we will address in this paper is the difference in the behavior in changes in the environment, as shown on the right side of Figure 1. Therefore, the methods used in these studies cannot be applied to this problem.

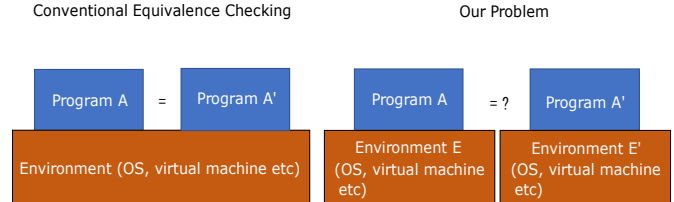


Figure 1: Differences from Our Previous Study

We did not find any research on the behavioral equivalence of Bash when the environment changes. Therefore, it would be useful for developers to provide a tool to verify whether a script behaves equally across operating systems.

In this paper, we propose a method to solve the problems that occur when updating the OS by verifying whether the behavior of Bash scripts is equivalent between two different version OSes and detecting the cause of the inequality. There are many distributions of Linux, but in this paper, we focus on CentOS, which has been used for business systems in companies.

### 1.2 The Approach

In the proposed method, we embed commands which generate execution log in the script file to be verified. The log generating commands(Loggers) are programs that log standard output and variable assignments. The embedded scripts are executed to generate and retrieve the logs. This execution process is performed on each two target OS built on VirtualBox[10], and the obtained logs are compared to detect differences in behavior. If the logs match, it is assumed that there are no changes in the behavior. If the logs don’t match, we present where the logs differ as the commands which is cause of the difference(hereinafter called “cause command”) to the developer.

We conducted two evaluations of the proposed method. In the first evaluation, we intentionally created commands with different behaviors and experimented to detect the difference in behaviors and to identify the cause. In the second evaluation, we experimented to detect the difference in the behavior of a script containing the command sudo, whose behavior was changed by updating the command in the past and to identify the cause of the change.

As a result of our evaluation experiments, we confirmed that we can detect differences in the behavior of scripts that contain commands with different behaviors. We found that it is possible to detect differences in the behavior of commands with small side-effects, for example, calculate and return arguments, but that it is difficult to directly identify the cause of

the differences in commands with side-effects, such as those that affect the environment variables.

In the following sections, Section 2 explain the motivating example, and in Section 3 we describe the proposed method. In Section 4 we present the results in the evaluation experiments and in Section 5 we discuss the results in the experiments. Finally, we conclude in Section 6.

## 2 THE MOTIVATION EXAMPLE

Listing 1: foo.bash

```
1 VAR=' baz '
2 export VAR
3 sudo bash bar.bash
```

Listing 2: bar.bash

```
1 echo ${VAR}
```

The two Bash scripts in the Listings 1 and 2, foo.bash and bar.bash, are examples of scripts whose behavior changes depending on past OS versions. In CentOS5 and later versions, the behavior of the sudo command in foo.bash has changed due to the upgrade of the command itself. Versions older than CentOS5 will output the string “baz,” while CentOS5 and versions later than CentOS5 will not output anything. That is because, in 1.6 and earlier versions of the sudo, the command was able to inherit environment variables when executed by sudo, however, in 1.7 and later versions, it is necessary to specify the ‘-E’ option to inherit environment variables. The environment variable is not referenced due to the change of the command specification on the newer version. This change causes the difference in behavior between versions older than CentOS5 using version 1.6 bash and later versions of CentOS.

## 3 THE PROPOSED METHOD

The proposed method generates execution logs of Bash scripts and judges the difference in behavior by comparing the logs. The execution log contains the standard output, the output between pipelines, and the variable assignments along with the executed commands. The stack trace information of the command is also stored to make it easier to identify the cause.

We create the tool based on the proposed method. The tool takes scripts and two target OS as input, embed logger commands, run embedded scripts, compare logs, and display results automatically.

### 3.1 Outline of the Proposed Method

The schematic diagram of the proposed method is shown in Figure 2. The input and output of the proposed method are shown below.

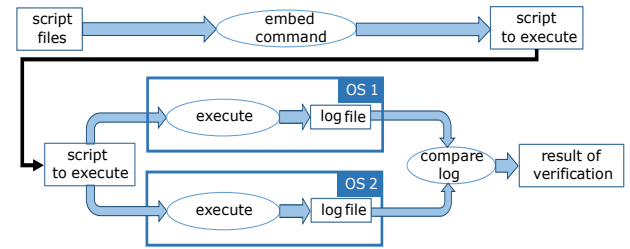


Figure 2: Diagram of the Proposed Methodology

- Input : Script file to be verified
- Output: verification results (the behavior is different and which command behaves differently)

The procedure of the proposed method is as follows.

(1) It embeds the Loggers into the Bash script given as input. The Loggers refer to a program that logs the standard output, output between pipelines, and variable assignment. The details of the Loggers are described in Section 3.3.

(2) It executes script embedded commands on each of the two operating systems.

(3) It compares the obtained logs and judges whether the behavior is different. If the behavior is different, the command that causes the difference and its log information is presented.

We have created a tool that automatically executes the above procedure.

### 3.2 Execution Log

The format of the execution log is shown in Listing 3.

Listing 3: Abstract of the Log

```
1 <log identifier>:<commands>, line: <
   lineno>, stack: <stack trace>
2 <log>
3 :<log identifier>
```

1. <log identifier> : “assignment” if the log is an assignment, “command” if the log is a command execution
2. <commands> : Executed commands
3. <line> : line number of executed command
4. <stack trace> : Stack trace information of the script and function when the command is executed.
5. <log> : Standard output or substitution log

### 3.3 Log Generating Commands(Loggers)

The following four Loggers are embedded in the script. The behavior of each is shown below.

- Standard output log command : Generate standard output log
- Assignment Log command : Generate Variable Assignment Log

- Stack push command : Record Stack trace information
- Stack pop command : Record Stack trace information

### 3.3.1 Standard Output Log Command

This command logs the standard output and the output between pipelines. It also takes two arguments, a command to be executed and its line and logs the string to be executed. The standard output log command is “stdout\_logger” in Listing 5. This logging command can generate the standard output log by pipelining this command to a line that does the normal standard output as shown in the examples in Listings 4 and 5.

Listing 4: Before Embedding Example for Standard Output

```
1 echo hello
```

Listing 5: After Embedding Example for Standard Output

```
1 echo hello | stdout_logger 'echo
  hello' 1
```

Running the script ex.bash that executes Listing 5 will generate the log shown in Listing 6.

Listing 6: Log Example for Standard Output

```
1 command:echo hello, line: 1, stack:
2 hello
3 :command
```

### 3.3.2 Assignment Log command

In this step, we generate log of assignments to variables. The assignment log command is “assign\_logger” in Listing 8. It takes the assignment command to be executed, the variable name, and the value of the variable as arguments and records them in the execution log. When the assignment command is taken as an argument, single quotes are escaped to avoid the expansion of variables. The Listings 7 and 8 show an example of before and after embedding the log.

Listing 7: Before Embedding Commands

```
1 VAR=' baz '
```

Listing 8: After Embedding Commands

```
1 VAR=' baz '
2 assign_logger 'var='\'' baz '\'' VAR
  $VAR 1
```

Running the script ex.bash that executes Listing 8 will generate the log shown in Listing 9.

Listing 9: Log Example for Assignment

```
1 assignment:VAR=' baz ', line: 1, stack
  :
2 VAR= baz
3 :assignment
```

### 3.3.3 Stack Push, Stack Pop Command

In order to identify the execution path of commands, command to stack push and pop records stack trace information into a text file. At the start of the script or function, their name is pushed. At the end of that, the pushed name is popped. In addition, the call command string is pushed just before the script or function call, and the pushed string is popped after the calling. Listings 12 and 13 shows the result of embedding for ex.bash and test.bash shown in Listings 10 and 11.

Listing 10: ex.bash Before Embedding Commands

```
1 test.bash 'test'
```

Listing 11: test.bash Before Embedding Commands

```
1 echo ${1}
```

Listing 12: ex.bash After Embedding Commands

```
1 push_stack ex.bash
2 push_stack 'test.bash '\'' test'\''
3 test.bash 'test'
4 pop_stack
5 pop_stack
```

Listing 13: test.bash After Embedding Commands

```
1 push_stack test.bash
2 echo ${1} | stdout_logger 'echo $
  {1}'
3 pop_stack
```

Listing 12 generate the log shown in Listing 14.

Listing 14: Log Example for Stack Push Command

```
1 command:echo ${1}, line: 1, stack:
  ex.bash->test.bash 'test'->test.
  bash
2 test
3 :command
```

The log “ex.bash->test.bash ‘test’->test.bash” which is the stack trace information following “stack:” in the log, indicates that the log was generated in test.bash called by test.bash ‘test’ from within ex.bash.

## 3.4 Comparing Logs

The tool compare 2 execution logs generated by running. Execution log contains following information.

- Executed command
- Stack trace information
- Log identifier of standard output or assignment
- Log of standard output or assignment

If no differences between the logs is detected, the tool judge that behavior is consistent. If not so, the tool judge that behavior is changed and display different logs as causes.

## 4 EVALUATING EXPERIMENT

In order to evaluate the proposed method, we conducted an evaluation experiment on the following two scripts to see if it is possible to detect and identify the cause of different behaviors.

1. A script that executes commands created to behave differently between operating systems.
2. Script containing the sudo command described in Section 2

### 4.1 Experiment 1

We prepared a command “sample” and a simple script to execute the command for each OS. The command “sample” takes two integer arguments and outputs the result of addition on CentOS4.6 and multiplication on CentOS8.2. The proposed method is applied to the scripts we created and conducted experiments to evaluate whether the tool can detect a difference in behavior between scripts with a different function, and whether “sample” can be identified as the causative command.

The script used for the experiment, “ex.bash,” is shown in Listing 15.

Listing 15: ex.bash

```
1 result=$(sample 2 3)
2 echo ${result}
```

Script after embedding the commands into ex.bash is shown in Listing 16.

Listing 16: ex.bash After Embedding Command

```
1 push_stack ex.bash
2 result=$( sample 2 3 | stdout_logger
   'sample 2 3' 1 )
3 assign_logger 'result=$( sample 2 3
   )' result $result 1
4 echo ${result} | stdout_logger 'echo
   ${result}' 2
5 pop_stack
```

Obtained logs by executing the above script on each target OSes are shown in Listings 17 and 18.

Listing 17: Log on CentOS4.6 in experiment 1

```
1 command: sample 2 3, line: 1, stack:
   ex.bash
2 5
3 :command
4
5 assignment:result=$( sample 2 3 ),
   line: 1, stack: ex.bash
6 result=5
7 :assignment
8
9 command: echo ${result}, line: 2,
   stack: ex.bash
```

```
10 5
11 :command
```

Listing 18: Log on CentOS8.2 in experiment 1

```
1 command: sample 2 3, line: 1, stack:
   ex.bash
2 6
3 :command
4
5 assignment:result=$( sample 2 3 ),
   line: 1, stack: ex.bash
6 result=6
7 :assignment
8
9 command: echo ${result}, line: 2,
   stack: ex.bash
10 6
11 :command
```

The results of the comparison of the two logs are shown in Figure 3.

Figure 3: Result of Experiment 1

Different logs are suggested. According to the results in Figure 3, differences in the script behavior were detected. The cause of the different behavior of the script “ex.bash” is the command “sample,” and the first presented log shows the command “sample 2 3.” Therefore, the difference in behavior was detected and the command that caused the difference was identified.

### 4.2 Experiment 2

The proposed method is applied to the scripts, which are shown in Listings 1 and 2, which are indicated in Section 2 and conducted experiments to evaluate whether the tool can detect differences in the behavior of the scripts and identify the command “sudo” as the cause of the difference.

The Bash scripts after embedding the commands into the scripts are shown in Listings 19 and 20.

Listing 19: foo.bash After Embedding Command

```

1 push_stack foo.bash
2 VAR='baz'
3 assign_logger 'VAR='\''baz'\'' VAR
  $VAR
4 export VAR | stdout_logger 'export
  VAR'
5 push_stack 'sudo bash ./bar.bash'
6 sudo bash ./bar.bash
7 pop_stack
8 pop_stack

```

Listing 20: bar.bash After Embedding Command

```

1 push_stack bar.bash
2 echo ${VAR} | stdout_logger 'echo ${
  VAR}'
3 pop_stack

```

The logs generated by executing the scripts shown in Listings 19 and 20 on each target OSes are shown in Listings 21 and 22.

Listing 21: Log on CentOS4.6 in experiment 2

```

1 assignment:VAR='baz', line: 1, stack
  : foo.bash
2 VAR=baz
3 :assignment
4
5 command:export VAR, line: 2, stack:
  foo.bash
6 :command
7
8 command:echo ${VAR}, line: 1, stack:
  foo.bash->sudo bash ./bar.bash->
  bar.bash
9 baz
10 :command

```

Listing 22: Log on CentOS8.2 in experiment 2

```

1 assignment:VAR='baz', line: 1, stack
  : foo.bash
2 VAR=baz
3 :assignment
4
5 command:export VAR, line: 2, stack:
  foo.bash
6 :command
7
8 command:echo ${VAR}, line: 1, stack:
  foo.bash->sudo bash ./bar.bash->
  bar.bash
9
10 :command

```

There is a difference in the 9th line of each log. CentOS4.6 outputs “baz,” but CentOS8.2 outputs a blank string. This is same as result mentioned in Section 2.

The results of the comparison of the logs are shown in Figure 4.

Figure 4: Result of Experiment 2

From the results in Figure 4, the difference in behavior is detected by comparison of the execution logs. However, the command “echo \$VAR” suggested as a cause is not true cause as described in Section 2. The true cause command “sudo” wasn’t identified as a cause of the difference in the behavior of the script. The direct cause, “sudo bash . /bar.bash” is included in the stack trace information “foo.bash->sudo bash . /bar.bash->bar.bash,” but it is difficult to identify the command “sudo” as the cause from the logs.

## 5 DISCUSSION

4.

In both Experiments 1 and 2, the difference in behavior between the script which executes commands with different behaviors is detected.

The cause command of the difference is identified in evaluation Experiment 1. In the experiment, differences in execution logs were detected in command substitution and assignment before they are outputted as standard output.

While, in evaluation Experiment 2, the wrong command was suggested as the cause. The reason is that the environment variables are not referenced in the script called by the “sudo” command in CentOS8.2, and the difference in behavior is surfaced by “echo \$VAR” which performs in standard output. In the case of changing the behavior of commands which has no standard output, differences in behavior are detected indirectly in standard output and assignment. This is because the proposed method generates execution logs with a focus on standard output and assignment. Therefore, it will be difficult to identify directly such commands as the cause. Similarly, it will be difficult to precisely identify the causative commands when the behavior of a command like “sudo,” which has a function to affect shared resources such as environment variables changed, i.e. a command with a strong side effect.

## 6 CONCLUSION

In this paper, we proposed a method to verify the behavior of shell scripts written in Bash before and after OS upgrade. By comparing the execution logs of shell scripts which are embedded Loggers, we conducted verification of the differences in behavior between the two OSes and identified the causes of the differences. It is confirmed that the difference in the behavior in standard output and assignment can be verified from evaluation experiments.

The shell script used in the evaluation experiment was composed of simple grammar without repetition and branching. The lexical analyzer defines only simple grammars and cannot support complex grammars. Therefore, it would be more



useful if more grammars could be supported by expanding the grammars.

Besides, since this method only collects logs that appear in the standard output or assignment, it is not possible to verify whether the behavior is changed when the behavior differs in areas that do not appear in the standard output or assignment, e.g. signal trapping. Thus, addressing these issues will be the main task in the future.

## ACKNOWLEDGEMENT

Part of this work is supported by fund from Mitsubishi Electric Corp.

We are deeply grateful to Mr. Satoshi Suda and Mr. Nobutoshi Todoroki, Mitsubishi Electric Corporation, for helpful discussion.

The research is also being partially conducted as Grant-in-Aid for Scientific Research A (19H01102) and C (21K11826).

## REFERENCES

- [1] “GNU Bash,” <https://www.gnu.org/software/bash/>
- [2] Jaechang Nam, Song Wang, Yuan Xi, and Lin Tan: “A bug finder refined by a large set of open-source projects,” *Information and Software Technology*, Vol.112, pp.164–175 (2019)
- [3] Sunghun Kim, Thomas Zimmermann, Kai Pan, and Emmet James Whitehead Jr.: “Automatic Identification of Bug-Introducing Changes,” 21st IEEE/ACM International Conference on Automated Software Engineering (ASE’06), pp.81-90 (2006)
- [4] Sokratis Tsakitsidis, Andriy Miranskyy, and Elie Mazzawi: “Towards Automated Performance Bug Identification in Python,” 2016 IEEE International Symposium on Software Reliability Engineering Workshops (ISSREW), pp.132-139 (2016)
- [5] Keigo Matsushita, Masaki Matsumoto, Kazuhiko Ohno, Takahiro Sasaki, Toshio Kondo, and Hiroshi Nakashima: “A Debugging Method Based on Comparison of Execution Trace,” *Symposium on Advanced Computing Systems and Infrastructures (SACSIS)*, Vol.2011, pp.152-159 (2011) (in Japanese)
- [6] Kozo Okano, Rin Karashima, Satoshi Harauchi, and Shinpei Ogata: “Regression Verification for C Functions with Recursive Data Structure,” *International Journal of Informatics Society*, Vol.11, No.2, pp.107-115 (2019)
- [7] Kozo Okano, Satoshi Harauchi, Toshifusa Sekizawa, Shinpei Ogata, and Shin Nakajima: “Consistency Checking between Java Equals and hashCode Methods Using Software Analysis Workbench,” *IEICE Transactions on Information and Systems*, Vol.E102, No.8, pp.1419-1422 (2019)
- [8] Rin Karashima, Satoshi Harauchi, Shinpei Ogata, and Kozo Okano: “Proposal and evaluation for property verification for Java functions with recursive data structures by SAW,” *Proceedings of International Workshop on Informatics 2019 (IWIN2019)*, pp.155-162 (2019)
- [9] “BASH Debugger,” <http://bashdb.sourceforge.net/>
- [10] “Oracle VM VirtualBox,” <https://www.virtualbox.org/>
- [11] Cherry Oo and Hnin Min Oo: “Spectrum-Based Bug Localization of Real-World Java Bugs,” *International Conference on Software Engineering Research, Management and Applications*, pp.75-89 (2019)



Session 6:  
Agricultural IT  
( Chair: Kei Hiroi )



# Proposal of a small robot for agricultural observation

Kenji Terada\*

Masaki Endo\*, Takuo Kikuchi\*, Shigeyoshi Ohno\*

\* Division of Core, Polytechnic University, Japan  
{k-terada, endou, kikuchi, ohno}@uitec.ac.jp

**Abstract** – Smart agriculture using ICT is spreading. What is important in precision agriculture is how to observe the conditions of crops and the environment. The use of drones to observe the conditions of widespread agricultural land is being investigated, but it is physically impossible to look up at the crops from below. Additionally, it is difficult to observe details of soil under the crops. Therefore, we propose the development of a small robot that can look up at crops from below, look horizontally, and observe them. This small robot makes it possible to observe soil details.

**Keywords:** agriculture robot, IoT, sensing

## 1 INTRODUCTION

The Japanese agricultural sector is confronting difficulties of an aging population and a severe labor shortage. As the number of agricultural enterprises continues to decline in Japan, the average size of arable land per farmer is expected to increase [1]. Therefore great interest has arisen in smart agriculture systems. Farmers are required to use internet of things (IoT) and robot technology to acquire and analyze highly accurate data from numerous factors related to crop production to realize precision agriculture [2]. These demands have accelerated the development of machines that are useful on farms [3][4].

Particularly, AI-based systems are finding new value in agricultural management. Currently, a cloud service that uses sensors to measure environmental conditions to predict harvest and disease outbreaks is in practical use [5]. For example, services all over the world determine disease damage from images of plants taken by farmers with smartphones and other devices [6][7][8][9]. An inexpensive system that uses crop images to monitor crop growth will increase agricultural efficiency. It is also expected to reduce the risk of crop failure in areas where there are few skilled farmers. In another study, remote sensing technology using drones is used to check the growing conditions through images. With the advent of drones, it has become easier to create data related to farm growth management [10]. However, the estimation of growing conditions by aerial images taken by drones requires ground-truth: actual measurements. Moreover, aerial photography by drones is not a panacea. Observations of the ground are necessary. For example, disease detection from imaging requires visual evaluation such as observing a single leaf on the ground for early detection of diseases such as powdery mildew on the underside of leaves [11]. These observations are limited to those made from the air. Various pests that damage crops, such as lepidopteran pests, coleopterans, and spider mites, are parasites that feed on the underside of leaves [12]. If even one virus-

diseased plant is missed, the infection might spread to other crops on the same farm, leading to secondary damage. The amount of damage caused by such diseases is said to exceed 100 billion yen per year [13]. The virus carrier is an insect or fungus that parasitizes the underside of a leaf [14]. Such carriers are a persistent difficulty on large farms, where it is difficult to look around and in areas with severe weather.

For the reasons described above, this study proposes a small agricultural robot that observes the soil and looking-up images of crops on a farm as a complement to the agricultural management data obtained from drones and farmers. Chapter 2 describes related studies and the position of this study. Chapter 3 describes a prototype developed for this research. Chapter 4 describes results of prototype driving tests. Chapter 5 presents the conclusion.

## 2 RELATED WORK

For this study, the name of the small robot to be developed is the Field Scouting Robot (FSR). Robots of many types patrol farms [15][16][17][18]. Actually, FSR aims at achieving stable autonomous running on uneven terrain, even in a small size, to observe close to many crops. Most of the small robots are four-wheel drive robots, but FSR chose to use six in-wheel motors to achieve stable ground contact on uneven terrain, as a farming machine. The rocker bogie mechanism used in the Mars explorer is famous for the same idea. The rocker bogie mechanism has a structure that can apply equal load to each wheel, which enables stable operation, even on uneven terrain. However, because of the structure of rocker bogie mechanism, a large difference exists in running performance between the front and rear. Additionally, it has been pointed out that the number of parts increases with the demand for lower height; it is not a simple mechanism [19]. Therefore, the frame of the FSR uses the mechanism of the six-wheeled vehicle that is different from the rocker bogie mechanism. The rolling mechanism allows the FSR to maintain contact between the ground and the six wheels even when the ground is uneven on both sides. The other mechanism used for reference uses in-wheel motors that are driven separately. Steering is done by differential movement of the left and right wheels [20]. The FSR is able to roll, so the FSR maintains contact between the ground and the six wheels, even when the ground undulations differ in the left and right directions. In other words, in terms of design, the mechanism alone can follow the undulations of uneven terrain and allow all wheels to contact the ground. There is also little difference in the runnability of the front and rear. The focus is on complementarity in case of failure. The aim is to enable the vehicle to run even if one unit is defective. Motion control enables turning by outputting different outputs

to the front and rear left and right wheels. Thus the optimal output can be given to the wheels for the required trajectory.

Fig. 1 presents the FSR position in farm observations. The FSR is aimed at complementing the analysis data obtained by drones that make observations from the sky, thereby observing areas that are difficult for farmers to see and which require farmer to observe, such as the base of crops and the underside of leaves. FSR will be used to observe the crops near the farm with thermometers, hygrometers, and barometers. Therefore, the FSR will be equipped with a camera to observe crops in real time. The situation can be checked with a smartphone or other device by streaming and storing crop images. By looking up at the crop, the sky will occupy most of the background. It is possible to reduce the amount of background information unrelated to the plant. The problem with plant disease diagnosis is that operational performance cannot be ensured for images taken in a different environment because of overtraining that includes background information which occupies a larger area than that of the plant [21]. By reducing the number of useless background images, the analysis time can be reduced.

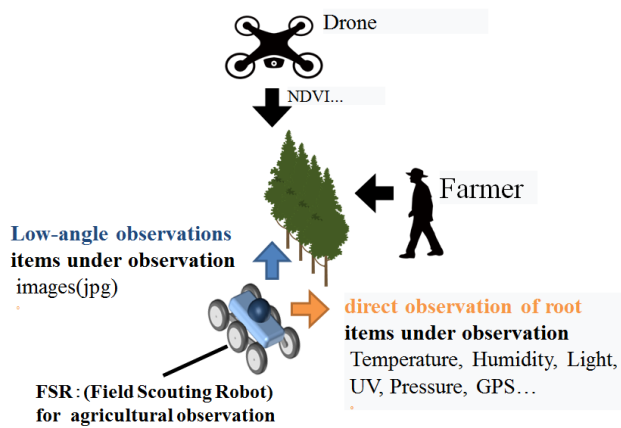


Figure 1: Positioning of FSR for agricultural observation.

Table 1: Specifications of FSR

FSR Length [mm]	350.0
FSR Width [mm]	294.0
FSR Height [mm]	310.0
Weight [kg]	7.2
Wheel width [mm]	65.0
Wheel weight [g]	760
Max. Power [kW]	$0.5^{*2} \times 6^{*1}$
Max. Velocity [km/h]	$30.3^{*2}$
Min. Velocity [km/h]	$2.4^{*2}$
Average uptime [min]	$63.3^{*3}$

\* 1 number of motors

\* 2 motor catalog data (over specification)

\* 3 no load

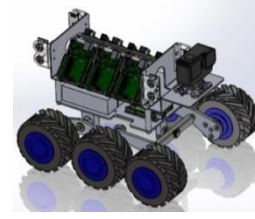


Figure 2: Prototype design of FSR.

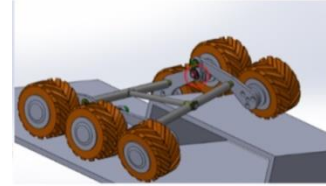


Figure 3: Design and simulation of FSR.

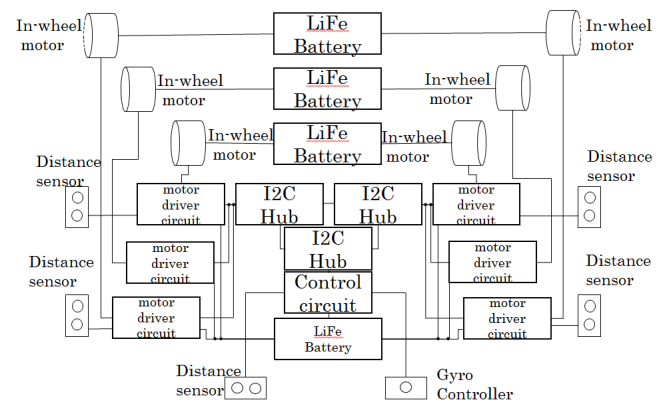


Figure 4: In-wheel motor control system of FSR.

### 3 PROTOTYPE ROBOT DEVELOPMENT

#### 3.1 Specifications of FSR

Fig. 2 and Table 1 present specifications of the FSR. Fruit trees, especially those that use leaves such as mulberry and tea, are harvested once a year. These crops need to be observed daily to maintain quality. The spacing between fruit trees in a small orchard is 1,000 mm [22]. The size of the FSR is aimed to be large enough to observe these crops. The in-wheel motor used for the FSR was selected from commercially available products that can be procured easily. We selected the in-wheel motor for the FSR from commercially available products that were easily procured. The two main points were that the motor should support only one side of the wheel (cantilever) and that it should have sufficient torque to tow its own weight to realize the mechanism used in the FSR. In-wheel motors for hard disks were unable to provide sufficient torque. In-wheel motors for scooters, of which there are many available on the market, have little cantilever. Those which could be procured were larger than expected. Eventually, we selected parts from a battery-powered E-Skateboard. The frame size was designed based on the selected in-wheel motor diameter. It was sufficiently large to observe an orchard farm of apples, mulberries, etc.

### 3.2 In-wheel Motor Control

Fig. 4 shows the control circuit of the in-wheel motor. The driver circuit that came with the in-wheel motor did not have the capability to connect to other electronic circuits. Therefore, We built a circuit using a brushless DC motor driver(TB6605FTG). The in-wheel motors on the market varied in terms of weight and the number of coil turns. In addition, individual speed control is necessary to achieve straight line operation [20]. A microcontroller for in-wheel motor control was provided for each in-wheel motor. The two microcontrollers for the front wheels and the two microcontrollers for the rear wheels are connected to an ultrasonic sensor [HC-SR04]. The ultrasonic sensor enables detection of obstacles on the side of the FSR. A control microcontroller was constructed to send rotation commands to the microcontroller which controls each in-wheel motor. To enable BLE communication with smartphone applications, as described in Chapter 3.4. The microcontroller for control is connected to the BLE module [BLE Serial3]. In addition, the control microcontroller is connected to the gyro sensor [CMPS12], which is located at the center of the FSR's chassis for angle control of turning and other operations. For communication between microcontrollers, microcontrollers use I2C communication. One battery was used for each of the two in-wheel motors to enable testing of the individual motor controls and flexible layout adjustments in the FSR. In addition, one battery was used for each of the seven microcontrollers. The battery material selected was Li-Fe, which is lighter and safer than Li-Po. Compared to Li-Po batteries, Li-Fe batteries are formed from materials that are extremely resistant to ignition, making them effective for robots with high vibration. The speed of the FSR for observation was set to 3.2 km/h, which is regarded as a slow human gait [23]. Because the selected in-wheel motor was not designed for low speed, the control microcomputer program had to reduce the responsiveness. Therefore, the program to lower the clock of the timer interrupt used for feedback control in the control microcomputer program made it possible to reduce the motor speed.

After assembling the FSR, a test course (Fig. 5) with a 30 mm step was used to verify that the FSR can maintain the same level of wheel contact as in the CAD simulation. As a result of visually checking the ground contact, FSR confirmed that the six wheels were installed on bumpy ground (Fig. 6).

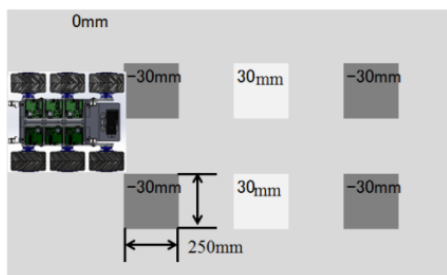


Figure 5: Ground contact performance on a test course with 30 mm bumps.



Figure 6: Driving of FSR with high grounding capacity.

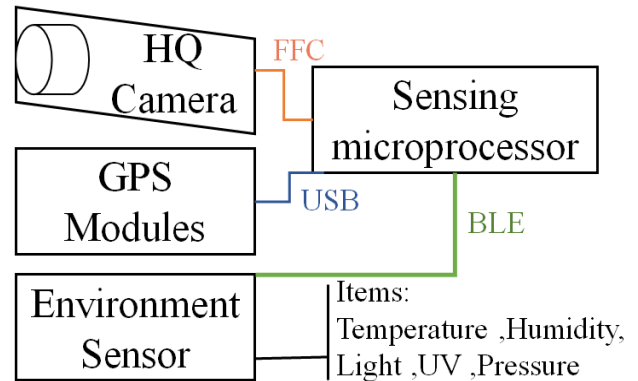


Figure 7: Log of environment sensor recording function of FSR.



Figure 8: Smartphone application to operate FSR.

### 3.3 Log of Environment Sensor Recording Functions

Fig. 7 shows the log of environment sensor recording functions of FSR. The FSR's sensing microprocessor uses a Raspberry pi4 B+. To obtain environmental information of date and time, temperature (degC), humidity (%), light (lx), U-V index, pressure (hPa), and noise (dB), FSR uses an environmental sensor (2jcie-bl01; Omron Corp.). The sensor can connect FSR's sensing microprocessor via BLE communication, so the sensing microprocessor location on FSR can be changed flexibly. To obtain images of the crops, the sensing processor is connected to a Raspberry Pi HQ Camera. In addition, to obtain the observation location coordinates, the sensing microprocessor is connected to a GPS Module (UI-

imate GPS Breakout - 66 channel w/ 10 Hz updates – Version 3; Adafruit Inds.). The program for the observation was programmed using Python. The measured environmental information is saved in a CSV file in the sensing microprocessor and in a DB (influxdb). The sensing microprocessor is installed with mjpeg-streamer to capture images and to distribute the images during observation for viewing on a web browser such as a smart phone. Using mjpeg-streamer, it is possible to distribute video and to acquire still images using the HTTP protocol. By observing log data gathered by FSR with the program, one can obtain environmental sensor data for crops.

### 3.4 Smartphone Application to Operate FSR

Fig. 8 depicts a screenshot of the smartphone application to operate FSR. An Android OS smartphone application controls the FSR. The camera image of the sensing microprocessor is projected at the top of the smartphone application screen. This screen is a web browser screen. The user taps the connect button to start communication with the FSR. After communication starts, the user can command the FSR to operate by pressing the respective arrow buttons. When the arrow button is released, the FSR stops. To experiment with turning as a prototype, pressing the automatic button will make the FSR turn when going straight, up to the length (m) entered.

## 4 EXPERIMENT

### 4.1 Mobility

Fig. 9 shows the movement of the FSR in a farm. To realize stable observation on the farm, FSR needs not only straight-line control but also control that can change the turning radius according to the size of the crop in the farm. The FSR is not equipped with a mechanism to change the wheel angle such as a constant velocity joint, so turning is achieved by the speed difference between the left and right wheels. Therefore, we analyzed the characteristics during turning by controlling the speed of the in-wheel motor.

Basic characteristics of the turning radius are described in

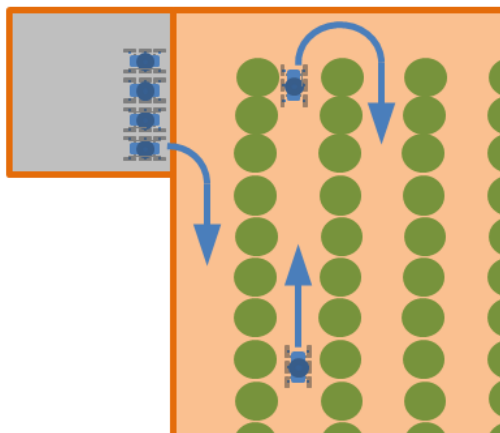


Figure 9: FSR in action on a farm

the six-wheeled vehicle model simulator described in one report [24] and the turning method for multi-wheeled vehicles in another report of the literature [25]. According to those reports of earlier studies [24] and [25], if the number of wheels of the driving unit increases and if the ground area becomes larger, then the resistance during turning increases and the turning radius becomes larger. For case of two wheels, the turning radius is the distance between the centers of the left and right tire widths when only one wheel is turned. The calculated value is 110 mm if we calculate the turning radius based on the FSR size. In reality, the turning radius will be larger because the ground resistance of FSR with six wheels is greater than that of two wheels.

To measure the turning radius repeatedly for the experiment, we used a board floor, which ensures more level ground than on soil, such as on a farm. We sent a left turn command from our smart phone application and drove until the FSR turned 90 degrees. The speed set for each wheel was  $100 \text{ min}^{-1}$ , which is the minimum speed at which the prototype can drive. The speed at which it stops was set to  $0 \text{ min}^{-1}$ . However, the wheels are not braked, so if an external force is applied, the wheels will turn. Table 2 shows the turning radius resulting from the difference in the way the left and right wheels turn. The measured values in Table 2 are the average values of 10 measurements in each case. The case of the spin turn caused by driving with all left wheels backward and all right wheels forward was excluded because it rotates around the chassis.

One can select a movement method that is useful for turning from the cases shown in Table 2. The method in case 1 is to operate only one wheel. The turning radius value of case 1 was 2,134.4 mm. The radius required for turning fruit trees varies depending on the farm. The general fruit tree spacing is 4,000 mm [26]. Case 1 can be found to be sufficiently operational as an orchard turning radius. However, the spacing between fruit trees in a small orchard is 1,000 mm [22], making it difficult to operate with the turning radius of the case 1 operation. Therefore, as shown for cases 2–7, we conceived a method of driving the left wheel, which has three wheels, backward when turning left as a method to reduce the turning radius. Cases 2–4 had a turning radius close to that of a super new land turn. The common feature was that two wheels on each side were driven backward. Therefore, we choose the rotation pattern to be used for turning from cases 5–7. Then, we measured the angle change during rotation in cases 5–7 patterns using the gyro-sensor attached to the FSR. Fig. 10 portrays a graph of the change in rotation angle for cases 5–7. In case 6, the return of rotation is large. It can be confirmed that it shakes during rotation. If the oscillation is large, then stable observation such as image recording by the camera cannot be expected. Therefore, it is necessary to select case 5 and case 7 for observation. For this study, case 7 was used to ascertain whether the turning radius increases because of the speed difference between the left and right wheels. Table 3 presents results obtained from increasing the rotation speed of the right wheel in the rotation pattern of case 7. Increasing the rotation speed of the left wheel caused a larger turning radius. Through these experiments, we were able to find a way to change the turning radius of the FSR.



Table 2: Turning radius caused by the difference in the way the left and right wheels turn (Fix the rotation speed of the left wheel)

	Pattern	Average turning radius [mm]
case 1	Right Wheels Front: forward Middle: forward Rear: forward Left Wheels Front: stop Middle: stop Rear: stop	2,134.4
case 2	Right Wheels Front: forward Middle: forward Rear: forward Left Wheels Front: stop Middle: back Rear: back	2.6 Near spin turn
case 3	Right Wheels Front: forward Middle: forward Rear: forward Left Wheels Front: back Middle: stop Rear: back	3.2 Near spin turn
case 4	Right Wheels Front: forward Middle: forward Rear: forward Left Wheels Front: back Middle: back Rear: stop	2.1 Near spin turn
case 5	Right Wheels Front: forward Middle: forward Rear: forward Left Wheels Front: stop Middle: back Rear: stop	145.3
case 6	Right Wheels Front: forward Middle: forward Rear: forward Left Wheels Front: stop Middle: stop Rear: back	457.7
case 7	Right Wheels Front: forward Middle: forward Rear: forward Left Wheels Front: back Middle: stop Rear: stop	121.0

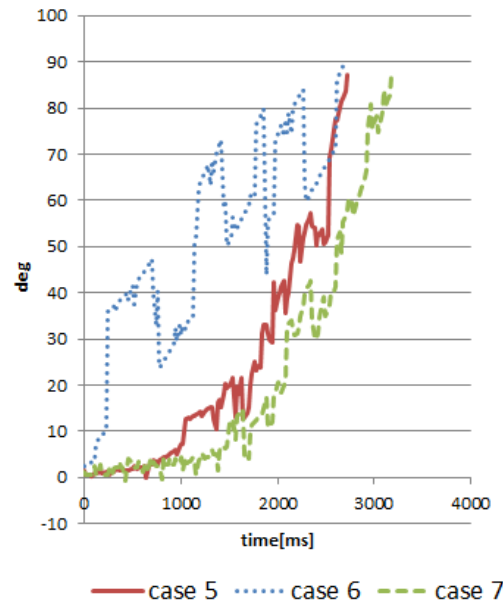












Figure 10: Change in rotation angle during FSR rotation.

Table 3: Turning radius caused by the difference in the way the left and right wheels turn (Fix the rotation speed of the right wheel)

	Pattern	Average turning radius [mm]
case 7-1	Right Wheels: forward (100 [min <sup>-1</sup> ]) Left Wheels Front: back (110 [min <sup>-1</sup> ]) Middle: stop Rear: stop	152.6
case 7-2	Right Wheels: forward (100 [min <sup>-1</sup> ]) Left Wheels Front: back (120 [min <sup>-1</sup> ]) Middle: stop Rear: stop	176.4

Table 4: Observation results for the mulberry farm

					
Observed image					
Pests and diseases by the system	powdery mildew	powdery mildew	aphid	aphid	yellow spot virus
Visual confirmation	un-observation	un-observation	observation	observation	observation

## 4.2 Observation of Leaf Undersides

Using sensors embedded in the FSR, we conducted an experiment to observe whether pests and diseases can be observed from observation results. Table 4 shows observation results in the mulberry farm. For diagnosis of diseases and insects, we used SCIBAI [8], a smart phone application that estimates diseases and pests from images. Mulberry

trees on the farm were photographed from above, and rows of trees without disease were selected. We drove the FSR along the side of the crops and looked up to observe. After sending the acquired images to SCIBAI for diagnosis of pests and diseases, we checked the location where we obtained powdery mildew detection results. However, no symptoms of powdery mildew were found. Powdery mildew causes white spots on the leaf surface. When the image was checked, we were able to observe that the leaves had spots through which light shined because of backlighting. We presume that these points were misdiagnosed as powdery mildew. We then obtained detection results of aphids from two locations. We were able to confirm the presence of aphids by checking the observed sites directly. Additionally, we observed yellowing virus disease from one location. Direct confirmation of the area in which the virus was observed revealed a yellow discoloration.

## 5 CONCLUSIONS

In this study, to gather environmental information related to the underside of crop leaves and the ground, we have produced a prototype of the FSR that can move autonomously, even on uneven terrain. After confirming that the FSR runs on uneven ground, we identified a method to find the turning radius through experimentation. Additionally, the observation function was found to allow FSR to obtain environmental values of the farm. The next theme is to develop a method for pest and disease prediction and detection based on observed images of pests and diseases, time, temperature, humidity, and other observed values obtained using detection technology. Furthermore, to enable autonomous operation, a patrol function based on longitude and latitude information from Camera, LiDAR and GPS is required. It must meet technical standards for automatic operation set by the Ministry of Agriculture, Forestry and Fisheries [27].

## REFERENCES

- [1] Statistics of Agriculture, Forestry and Fisheries, MAFF, pp. 11-7 (2020).
- [2] E. Shibusawa, Precision Agriculture, Asakura Publishing Co., Ltd.
- [3] J. Champ, A. Mora-Fallas, H. Goëau, E. Mata-Montero, P. Bonnet, Instance segmentation for the fine detection of crop and weed plants by precision agricultural robots, Applications in Plant Sciences, Vol. 8, No. 7 (2020).
- [4] O. Spykman, A. Gabriel, M. Ptacek, M. Gandorfera, Farmers' perspectives on field crop robots – Evidence from Bavaria, Germany, <<https://www.sciencedirect.com/science/article/pii/S0168169921001939>>, [referred May 2021]
- [5] “midori cloud (in Japanese)”, <<https://info.midori-cloud.net/vision/>>, referred May 2021]
- [6] H. Iyatomi, Trends and Challenges of Automatic Diagnosis Techniques for Plant Diseases, Japanese Neural Network Society Journal. Vol. 26, No. 4, pp. 123–144 (2019).
- [7] D. Mhlanga, Artificial Intelligence (AI) and Poverty Reduction in the Fourth Industrial Revolution (4IR). Preprints 2020, 2020090362 (doi: 10.20 944/ preprints-202009.0362.v1). (2020).
- [8] “SCIBAI (in Japanese)”, <<https://www.mirai-scienc.com>>, [referred May.2021]
- [9] “xarvio”, <<https://www.xarvio.com/be/nl.html>>, [referred June 2020]
- [10] R. Sugiura, Remote sensing for large-scale field information using drone imagery, Research Center for Agricultural Information Technology, pp. 184–187 (2020).
- [11] S. Yamamoto, Fundamental Study on the Collection of Big Data for the Agricultural Community Using Machine Vision: Exploring a Potential Technology to Complement Drone Information for Monitoring Plants, Akita Prefectural University Web Journal B, pp. 85–90 (2018).
- [12] Ministry of Agriculture, Forestry and Fisheries, pest control, <[https://www.maff.go.jp/j/seisan/kankyo/hozen\\_type/h\\_sehi\\_kizyun/attach/pdf/aki3-15.pdf](https://www.maff.go.jp/j/seisan/kankyo/hozen_type/h_sehi_kizyun/attach/pdf/aki3-15.pdf)>, [referred May.2021]
- [13] K. Oshima, Plant potyvirus evolution the survey of the genetic structure of populations, Virus, Vol. 2, pp. 151–160 (2012).
- [14] Japan Plant Protection Association, “Disease and Pest Control (in Japanese)”, <[https://jppa.or.jp/wpsite/wp-content/uploads/byougaichu/byougai\\_boujo.pdf](https://jppa.or.jp/wpsite/wp-content/uploads/byougaichu/byougai_boujo.pdf)>
- [15] M. Arun et al., “Smart Agriculture Robot”, International Journal of Pure and Applied Mathematics, Vol. 119, No. 15, pp. 1901–1906 (2018).
- [16] “avo”, <<https://www.ecorobotix.com/en/avo-autonomous-robot-weeder/>>, [referred Mar.2021]
- [17] “Weed Whacker”, <<https://www.odd.bot/>>, [referred Feb.2021]
- [18] “DICK”, <<https://www.smallrobotcompany.com/meet-the-robots>>, [referred Feb.2021]
- [19] A. Schafer et al., “Robot Mobility Concepts for Extraterrestrial Surface Exploration”, IEEE, Aerospace Conference, pp. 1–12 (2008).
- [20] Y. Hirota, Electric vehicle engineering (in Japanese), Morikita Publishing Co., Ltd. pp. 60–66 (2017).
- [21] T. Hiroki et al., “Diagnosis of multiple cucumber infections with convolutional neural networks”, IEEE Proc. AIPR, pp. 104 (2018).
- [22] The National Agriculture and Food Research Organization, “Small Fruit Tree Manual (in Japanese)”, (2018)
- [23] Ministry of Health, Labour and Welfare, “Physical activity standards for health in 2013”, pp. 51 (2013).
- [24] M.H. Prio, F. Rios, “Kinematic Modeling of a Six Wheeled Differential Drive Intelligent Robot and Potential Field Method to Attain Obstacle Avoidance Capability,” 2019 Southeast Con., pp. 1–4 (2019).
- [25] N. Ito, N. Iguchi, Steerability Control of Multi-Powered Wheel Vehicle, JSAM Journal, Vol. 50, No. 1, pp. 11–18 (1988).
- [26] “How to plant fruit tree seedlings (in Japanese)”, <<https://minorasu.basf.co.jp/80083>>, [referred Mar. 2021]
- [27] The National Agriculture and Food Research Organization, “Small Fruit Tree Manual (in Japanese)”, (2018)



# Implement of low cost based IoT system in the paddy field to labor-saving and feasible study in Otari village, Japan

Kazuma NISHIGAKI\*, Kanae MATSUI\*\*

\*Graduate School of Informatics, Tokyo Denki University, Japan

\*\*School of System Design and Technology, Tokyo Denki University, Japan  
{23jkm23@ms, matsui@mail}.dendai.ac.jp

**Abstract** - Rice terraces, which are one of the factors that form the traditional landscape of Japan, exist in the hilly and mountainous areas. Due to their topographical features and the aging of agricultural workers, it is difficult for rice cultivation to keep for a long time. Therefore, we focused on water management, which occupies about 30% of the working time in rice cultivation and proposed a paddy field water level monitoring system aiming at labor saving using Sigfox communication. We implemented a system that can be used by elderly farmers, and conducted an experiment in Otari Village, Nagano Prefecture to verify the availability of the proposed system.

**Keywords:** IoT application, Agriculture, Paddy field, Sigfox, Labor saving

## 1 INTRODUCTION

In the field of agricultural IoT, technological innovations in networks, hardware, and software have been progressing. In addition, studies and practical applications for the purpose of labor saving and automation have been promoted. In this field, the investment amount and the requirements for technology differ greatly between full-time farmers and side business farmers. In addition, since the applicable technology differs depending on the breeding environment, a system that meets the requirements as the agricultural system is required.

With this background, this study focused on labor saving in water management for part-time farmers in terraced rice fields. Rice terraces are not only valuable as a landscape of Japanese “satoyama”, but also an important resource in the region because they function as disaster prevention such as taking irrigated land. However, they are difficult to pave and are not flat. From these reasons, the rice terraces have been shrunk, and such a phenomenon occurs in the mountainous areas with rice terraces.

Therefore, we propose an IoT-based water level monitoring system that can be used by elderly farmers at low cost with the aim of saving labor in rice terraces. Water management accounts for 30% of the agricultural work during rice cultivation, and we aimed to relieve the burden by saving labor.

## 2 RELATED WORK

This chapter describes water management systems and applications for rice cultivation in the field of agricultural IoT.

### 2.1 PaddyWatch

This product is a system that automatically measures the water level and temperature required for paddy rice cultivation, which also has a water level sensor in paddy fields sold by Vegitalia Co., Ltd. [2]. This product aims to reduce the number of water patrols and the time required for patrols by farmers by providing the water level status on a web application. In addition, it has the functions necessary for proper water management, and additional functions for data analysis.

### 2.2 paditch gate 02+

This product realizes real-time remote control of sluice gate management as one of water managements provided by Enowa Co., Ltd. [3]. The product has a water level adjustment function, and users can select a time in advance to close a water gate for paddy fields. Also, the product can open and close the floodgate automatically. The sensor unit measures the water level and temperature of the paddy field. In water management, determination and actions such as whether to open the floodgate from data such as water level and temperature and add new water to the paddy field to raise the water level.

### 2.3 Position of this research

The agricultural population in rice cultivation is reducing, and the population is aging [4]. As mentioned above, there are many terraced rice fields in the mountainous areas, and the area of one paddy field is not large as an agricultural land, so a large-scale yield cannot be expected. In addition, as the location may not face the maintained roadway, there are disadvantages in terms of location such as difficulty in inserting agricultural machinery [5].

However, terraced rice fields have various functions, and considering the sustainability from the viewpoint of land use in mountainous areas, it is considered that it is an important measure to incorporate labor saving and automation technology in order to protect the terraced rice fields. In order to keep the rice terraces as cultivated land, we thought that it was necessary to pay attention to the bearers and to have a system that matches the attributes of the agricultural population. In most agricultural IoT systems, the information providing is from a web application. However, for elderly people, web applications are often not tools that they use on a daily basis. In this study, we customized and

provided the information by reflecting the contents of the information provided in advance.

### 3 PROPOSED SYSTEM

The proposed system considers the network environment of Otari Village, Kitaazumi District, Nagano Prefecture, which cooperates proof-of-concept of this study, and the characteristics of farmers in rice terraces. The outline of the system is shown below, and then the details are described for each network, hardware, and software.

#### 3.1 System overview

Figure 1 illustrates the overview of this study. In the following, the technologies for configuring the system are shown separately for networks, hardware, and software. The paddy field water level monitoring system shows a schematic diagram of this system in Fig. 1 for terraced rice fields in mountainous areas.

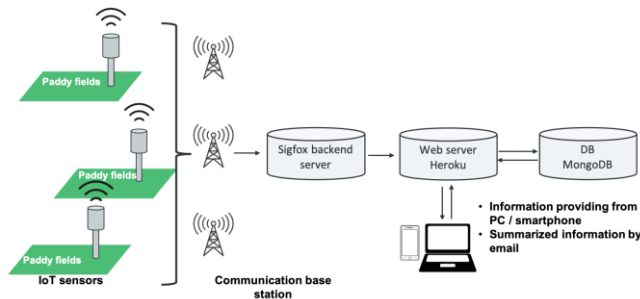


Figure 1 Overview.

The flow from data measurement to information provision shown in the figure is explained. First, the device that measures the water level and the other data in paddy fields has a built-in communication module with Sigfox and sends data at specified time intervals. Then, the data is sent to the Sigfox backend server, which sends it to the data server, which we developed in MongoDB. This data server communicates with the web application server prepared for the web application for users, provides a web page that provides data visualization to each user. Hence, measurement data as summarized information is sent to the specified e-mail address.

The details of the network, hardware, and software, which are the technologies for configuring this system, are described below.

#### 3.2 Network

As the network, we adopted Sigfox, which has been prepared by the local governments of Otari Village. Sigfox is one of the LPWA (Low Power Wide Area) standards, and is a global IoT network featuring low cost, low power consumption, and long-distance transmission [6]. Using the frequency band around 920MHz of this network, the speed of uplink communication from the terminal to the base station is about 100bps, and that of downlink communication is about 600bps. In addition, the maximum data capacity that can be transmitted in one communication

is 12 bytes, and the upper limit of the number of communications per day is set, and there is a limit of 140 times for uplink and 4 times for downlink.

The details of data selection will be described in the measurement section. This time, the data in Table 1 was acquired and the payload of the data sent from the hardware to the Sigfox Backend Server was in the format shown in Table 2.

Table 1 A list of collected data

Data	Contents
Water level 1	5 levels (minimum value 3 cm to maximum value 5 cm)
Water level 2	0.5 cm (maximum 12 cm)
water temperature	Paddy water temperature (°C)
temperature	Temperature (°C) about 1.2m from the ground
Humidity	Humidity on the board (%)
Atmospheric pressure	(hPa)
Substrate temperature	Substrate temperature (°C)
Illuminance	0-255 step illuminance
Internal operating voltage	(V)

Table 2 Detailed information of payload design

Byte index	Contents
0	Outdoor temperature
1	
2	Water temperature
3	Board temperature
4	Luminance
5	Voltage
6	Air pressure
7	
8	Humidity
9	Water level
10	
11	Empty

#### 3.1 Hardware

This section describes the network-compatible water level measurement device installed in paddy fields, including the unit of measurement, transmission, and power supply. Figure 2 shows a block diagram showing the connections between each part and the board.

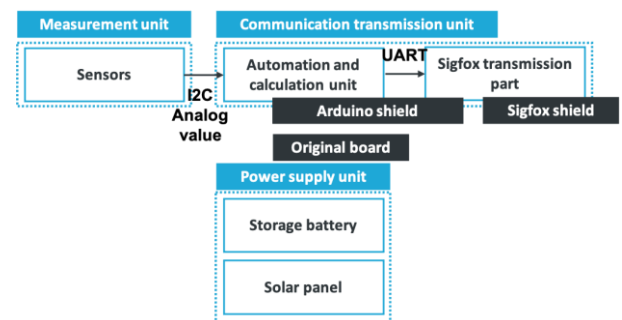


Figure 2 Block diagram

The hardware is divided into a data communication unit, a measurement unit, and a power unit, and the measurement unit mainly measures water level data and environmental data such as temperature and humidity.

### Communication department

The details of the communication unit are described. The communication unit consists of an Arduino that performs control and calculation as the main board, a Sigfox shield that performs transmission, and an original board that controls the power supply unit. The role of the communication unit is from the measurement unit to analog values and, the purpose is to send 9 types of data (Table 1) collected by I2C communication to the Sigfox Backend Server. A Sigfox module that can be connected to the Sigfox network is connected to the transmission unit, and data is transmitted to the data Sigfox Backend Server at the transmission timing and transmission interval specified by the control unit. Table 3 shows the details of the measurement items and the details of each measurement item of the Sigfox payload.

Table 3 Detailed information of collected data

Byte Index	Type	Data	Details	Error	Value
0	Int16	temperature	The temperature of the sensor installed at a position about 1.2 m above the paddy field	-1270	275 / 27.5 °C
1	Int8	water temperature	Paddy water temperature Decimal point truncation	-127	27 / 27 °C
2	Int8	Substrate temperature	Substrate temperature Decimal point truncation	127	27 / 27 °C
3	Int8	Lumina nce	0 - 255	-	240 / 240
4	Uint8	Internal operating voltage	Operating voltage of the main IC inside the board, rounded down to the first decimal place	-	50 / 5.0 V
5	Uint8	Barometric pressure	Atmospheric pressure on the board, Rounded down to the first decimal place	65535	10132 / 1013.2 hPa
6 <sup>*7</sup>	Uint16	Humidity	Humidity on the board, Decimal point truncation	255	42 / 42 %
8	Uint8	Water level	6 levels of water level measured with a float switch	255	1 / 1
9	Uint8	Water level	25 levels of water volume measured by water sensor every 5 mm	255	3 / 1.5 cm

10	Uint8	Substrate temperature	The temperature of the sensor installed at a position about 1.2 m above the paddy field	-	-
11	empty	unused	-	-	-

The data transmission timing is summarized below.

- When the water level changes: Measured every 10 minutes, and data is transmitted when the amount of water level change from the previous transmission value changes by 1 cm or more for the water level sensor and 1 step or more for the float switch.
- Regular transmission: Every 30 minutes

### Measurement unit

In this system, the water level and water temperature required for water management in paddy fields as collecting important measurement data. We used two devices with different measurement methods. One is a float switch type water level measurement sensor (hereinafter referred to as a float sensor) that determines the water level position by reacting the reed switch by raising and lowering the float. The other is a floatless switch type water level measurement sensor that measures the resistance value (voltage) between two poles and determines the presence or absence of water [7].

The former was intended to be adopted in this system because it can be produced with a relatively inexpensive configuration, but the measurement interval depends on the size of the module. Therefore, when designing from an existing module, the minimum measurement interval was 3 cm. In general, in paddy field management of individual farmers, we expected that the resolution would be sufficient and proceeded with the production. For the latter, the resolution is sufficient for water level measurement with a float sensor, a water level sensing sensor (ASZ-M0917 [8] made by Aszac Co., Ltd.) that realizes a measurement interval of 0.5 cm. Two types of sensors were used together so that.

In this study, we designed a float sensor type water level measurement sensor. Float sensors were installed at intervals of about 3 cm, and fine adjustments were made at the time of installation where the threshold value was requested. As mentioned above, the minimum value can be adjusted to about 3 cm, but it is possible to register the desired interval of the user in the system. If the interval is 5 cm, the measured quantity is 0 to 20 cm. The float sensor of S1 in the figure shows the ground side, and the resistance value (SUM R1 to R5) changes by grounding in order from S1, and the design enables identification by the input voltage.

Table 4 A table of Analog voltage threshold

	Measured value(mV)	Theoretical value(mv)	THRSD+	THRSD-
FREE	5000	5000	0	418
S1	4169	4165	418	83
S2	3998	4000	83	125

S3	3758	3750	125	208
S4	3450	3335	208	418
S5	2917	2500	418	2500
FREE	5000	5000	0	418

We collected not only the water level but also environmental data such as water temperature, outside air temperature, humidity, and atmospheric pressure, as well as state data inside the device (Table 1). The data are necessary for rice cultivation and to utilize the internal state data of the device. A sensor (DS18B20) was used for water temperature and air temperature, and a sensor mounted on Sigfox's transmission module Una Shield (V2S2) was used for collecting data of humidity, atmospheric pressure, and substrate temperature. The illuminance was converted from the input terminal of Arduino, which is capable of analog input of 0 to 5V, into 256 steps of 0 to 255 using a diode. The conversion is an Arduino specification, and the resolution is 19.5 mV.

Regarding the illuminance, it is assumed that the timing of sunrise and sunset, the relative shadow when compared with other days will be used as a reference. Hence, it is not assumed that it will be converted into a physical quantity that indicates the actual brightness of light. The internal operating voltage is the voltage that operates on the main IC in the Arduino. A voltage of 9 to 24V is input from the battery to the Arduino, and the voltage is internally stepped down to 5V. Since the value is the voltage collected in the main IC, 5V is measured in principle.

### Power supply part

About the power supply unit, it uses solar panels to generate electricity and a storage battery installed inside the device, and it rains for 2 to 3 days after being fully charged. In addition, the solar panel should be installed at a position different from the sensor device. Table 4 shows the electrical characteristics of the solar panel used this time. Table 5 shows the electrical characteristics of the device. Figure 6 shows the appearance of the hardware.

Table 5 Characteristics of solar panel

Item	Detailed information
Normal state	12W
Maximum output power	21.8V
Open circuit voltage	0.73A
Short-circuit current	17.4V
Maximum output power voltage	0.69A
Maximum load current	-

Table 6 Characteristics of electrical device

Item	Use
Storage battery capacity	14000 (mAh)
Input voltage	9~24 (V)
Operating voltage	9 (V)
power consumption	Standby: Approximately 0.36 (W) When sending: Approximately 0.9 (W)

Operating temperature	0 - 40 (°C)
Storage temperature	-40 - 85 (°C)

## 3.2 Software

The above-mentioned hardware in a paddy field collects data for paddy field monitoring such as water level, visualizes it with a device. In terms of software, it has three functions as a web application. Such as a smartphone or PC and provides an environment where the status of the paddy field can be confirmed even from a remote location. Following points were developed for the functions of web application.

1. Periodically collect measurement data from Sigfox backend server
2. Display the measurement data on the web page in user screens
3. Send the latest measurement data to the specified email address at the specified transmission interval

The function of item (1) is to store data from Sigfox backend server in a data server prepared in advance. For that we developed webhooks on the Sigfox backend server side and store it in the prepared database. "Webhook" is a mechanism to notify an external service by HTTP when an event is executed in a web application. When new data is stored in Sigfox backend server, the data is sent to the database at the same time. As item (2) under this function, the DB acquires the latest data and reflects the latest data on the web page from the linked web application server. In addition, as item (3), regular sending to e-mail will be carried out.

We built the web application server in Heroku, which is a cloud application platform, using Node.js, which is a JavaScript language. In addition, MongoDB, a document-oriented database, was used as a DB for storing measurement data, and data was linked between Heroku and this database using socket communication. In Heroku, we built a front side that presents Web pages built using the JavaScript framework Vue.js and the CSS framework Bootstrap. Figure 3 is implemented for the above functions. Table 7 shows the development environment.

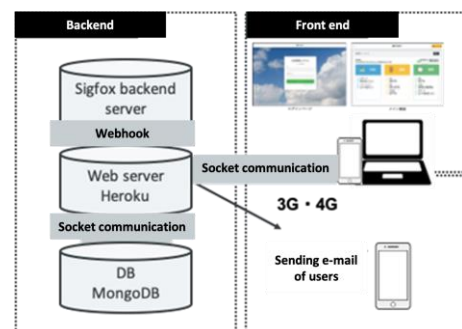


Figure 3 Overview of proposed system.

Table 7 Development environment of the software

Object	Purpose
Macbook	Development PC
Heroku	Web application server
MongoDB	Data server
Node.js	Backend and frontend development languages
Vue.js	Front-end development language
Bootstrap	CSS framework for front-end development

Figure 4 shows a screenshot of the actual web page. On the main screen, users can browse data by water level, temperature, environment and category, and the gauge on the right shows the time until new data is sent to Sigfox backend server, and the latest data is reflected on this page.



Figure 4 Screen shots of login and main web pages

In addition, this web page design is created from interviews conducted in advance with the experiment participants who are paddy fields owners. The participants have the Internet environment, have no resistance to digital devices such as personal computers, smartphones, and tablets, and are willing to use web applications, but select the types of data to be provided (mainly water level). The water temperature is easily noticeable, and they requested the push-type data provision by e-mail notification. Therefore, we implemented a simple design as shown in Fig. 8 and an email notification function.

#### E-mail notification

This function inserts the latest data in the database into a pre-made fixed phrase at the time specified by the user and sends it to the specified e-mail address. This function is executed in Heroku and received by the user through a wide area network such as 3G / 4G.

## 4 EXPERIMENT

This chapter describes an experiment for the proposed system. The experiment was conducted in Otari Village, Kitaazumi District, Nagano Prefecture for about two months from July 12 to September 20, 2019.

The experiment participants were three paddy field owners, and devices were installed in each of the three paddy fields, and a total of nine devices were operated. Before and after that, we conducted hearings to collect information related to this system. The information on collaborators, including geographical information, is described below.

## 4.1 Experimental participants

This section provides detailed information on the collaborators. Table 8 summarizes the information on collaborators. In order to keep personal information confidential, each participant was numbered by number.

Table 8 Information of experiment participants

Number	District	Age	Experiment target paddy field
N	Fukawara	70s	Three
C	Kurokawa	40s	Three
S	Mushio	60s	Three

In addition, Figure 9 illustrates the overall positional relationship. From the north, there are Fukahara, Mushio, and Kurokawa districts. Figure 10 shows the rice terraces in Otari Village.

## 5 EVALUATION

This chapter describes the evaluations collected from the experiments in networks, hardware, and software.

### 5.1 Network

This section summarizes the radio wave conditions in each area created based on the RSSI values stored in the Sigfox backend server. Three Sigfox base stations are installed in Kotanimura, and the identification IDs of the base stations are 65B2, 65A3, and 6595, respectively. Each base station is installed near the center of the identification ID circle shown in Figure 11.

Although Sigfox has a domestic population coverage rate of 95% in Japan [1], a base station was set up because there are areas where radio waves do not reach in Otari village, where is a mountainous area. Since this system was the first outdoor IoT system installed in the village, this analysis was conducted with the aim of confirming whether the network is working properly. It can be seen that the frequency of transmission to the base station of 65B2, which is the position shown in the figure. The same analysis was performed in the other two districts, and it was found that the frequency of transmission to neighboring base stations was high.

From this result, it was judged that the network was stable and did not lead to system operation problems such as data loss.

### 5.2 Hardware

This section mainly describes the measurement results of water level sensor data. Figure 12 illustrates all the data measured from the float switch and the water sensor from July 11, 2019, to September 19, 2019 (70 days) in Kurokawa district C-1. The horizontal axis represents the date, and the vertical axis represents the water level (0 to 15 cm). The amount of water in C-1 was relatively small, and the average value was close to 0. Comparing the value of the

float sensor with the value of the water sensor as a reference, it can be seen that the value fluctuates according to the value of the water sensor. However, from 8/5 to 8/7, 8/11, and 8/13, there are some places where the float sensor does not move.

### 5.3 Software

Since stable system operation was observed in the software, the results will be described in the interview based on the questionnaire conducted face-to-face in this evaluation. As items, the data obtained from this system and the usability of the Web application are described.

#### Data

The three collaborators browsed the Web application more than three times a day, and were highly interested in the measurement data. The most interesting data were the water level, followed by air temperature and water temperature. Since the data on air temperature and water temperature could not be confirmed so far, there were many opinions that they were wondering how they are related to the growth of rice. On the contrary, it turned out that the atmospheric pressure and the illuminance were not very helpful.

#### Web application usability

The convenience of the web application was highly evaluated, and it can be said that the usability was evaluated. However, there is a good opinion that necessary data such as the display of the maximum and minimum temperatures of the day can be displayed, and it can be said that the flexibility of the data display function is a future issue. In addition, the method of browsing Web applications differed depending on the personal computer, tablet, smartphone and the collaborators. Therefore, there was an opinion that the display was small, and it was considered necessary to make the font size customizable.

The purpose of this system is to reduce the number of patrols of paddy fields in water management by introducing the system, that is, to save labor. From the interviews with the collaborators, many people said that it was a reference for the patrol of the paddy field, and a certain evaluation was obtained for the efficiency of the patrol order of the paddy field and the labor saving of the overall number of patrols. Although there was not much change in the number of patrols per day, there was an opinion that it was reassuring to be able to make decisions on patrols in rainy weather and confirm information from a distance.

### 5.4 Overall

Since the purpose of this system is to save labor, we expected to reduce the number of patrols of paddy fields, but although it was helpful, the number of patrols of paddy fields did not decrease significantly due to the provision of information in this system. After all, there were many opinions that they wanted to grasp the rice growing situation by looking at the state of the paddy field, and it was difficult to provide an effect that exceeds the visual inspection of the rice growing situation from the provision of information.

## ACKNOWLEDGMENTS

We would like to express our deep gratitude to everyone at Otari Village Hall in Kitaazumi District, Nagano Prefecture, and KCCS Mobile Engineering Co., Ltd. for their great cooperation in this study.

## REFERENCES

- [1] Akira Ushijima, Masato Nakazono, Changes in farming environment and rice terrace conservation in small-scale villages in mountainous areas. Architectural Institute of Japan Technical Report Collection. 2017, vol. 23, no.55, p. 979-984.
- [2] Japan Cabinet Office, "Office products". URL: <https://office.microsoft.com/en-us/products>, Japanese only (2021-06-12).
- [3] "paddich". URL: <https://paditch.com/>, (2022-01-27).
- [4] "Basic policy on promotion of terraced rice fields". URL: [https://www.kantei.go.jp/jp/singi/tiiki/tanada/pdf/tanada\\_kihon\\_housin.pdf](https://www.kantei.go.jp/jp/singi/tiiki/tanada/pdf/tanada_kihon_housin.pdf), Japanese only (2022-01-27).
- [5] "Promotion of the terraced rice fields". URL: <https://www.maff.go.jp/j/nousin/tanada/tanada.html>, Japanese only (2021-06-12).
- [6] ZUNIGA, J. et al. "Sigfox system description," LPWAN @ IETF97, Nov. 14th, 2016.
- [7] Tung Ta Duc, Masaru Mizoguchi, Yoshihiro Kawahara, Tohru Asami, "Error Reduction in Capacitive Based Water Level Sensor for Paddy Field," Shingaku Sodai 2016, B-18-16, March 2016.
- [8] "Water level sensor (water level detection / detection / measurement)". URL: <http://www.asuzac-pd.jp/seihin/mizumisensa.htm>, Japanese only (2022-01-27).

# A Feature Generation Method for Plant Growth Prediction Using Random Forests

Yosuke Asada<sup>†</sup>, Hiromi Hanada\*, Takuya Yoshihiro<sup>‡</sup>

<sup>†</sup>Graduate School of Systems Engineering, Wakayama University, Japan

\* Horticultural Experiment Center, Wakayama Agricultural Experiment Station, Japan

<sup>‡</sup>Faculty of Systems Engineering, Wakayama University, 930 Sakaedani, Wakayama, 640-8510, Japan

<sup>†</sup>s226007@wakayama-u.ac.jp, <sup>‡</sup>tac@sys.wakayama-u.ac.jp

**Abstract** - The demand for agricultural crops fluctuates depending on the time of year and economic conditions. Therefore, it is important to respond to the market demand in a timely manner by controlling the harvest timing of agricultural crops. One example of a harvest time control method currently in use is the integrated temperature control. However, this empirical method is not based on any scientific ground, and the predicted harvest time has considerable errors. In addition, factors other than temperature, such as humidity, are important factors to determine the amount of growth, but are not considered. The purpose of this study is to predict the amount of growth using sensor data as a basic technology to control the harvest time more accurately. By predicting the growth of plant stems, it is thought possible to find factors in the environment such as temperature and humidity that are necessary for harvest time control. However, to make highly accurate predictions, the training data must contain environmental conditions and patterns over the past few days that promote plant growth. Therefore, in this study, we propose a feature generation method using agricultural sensor measurements over the past few days for predicting the amount of plant stem growth. The proposed method creates features by using sensor measurements over the past few days with respect to the predicted date. Plant growth does not depend on the response of plants on only one day, but depends on the changes in the environment over the past few days. Therefore, we attempted to predict the amount of daily growth by designing machine-learning features that reflect various conditions and patterns of past measurements. In the evaluation, the features generated from the sensor data using the proposed method were applied to the random forest method to evaluate the prediction accuracy of the amount of *Lisianthus* stem growth.

**Keywords:** Generating feature, Machine learning, Random Forest, Method for controlling the harvest time, IoT

## 1 INTRODUCTION

The demand for agricultural products differs depending on the time of year. It is required to respond to the market demand in a timely manner by controlling the harvest time of crops. One example of a method to control the harvest time is the integrated temperature control, which controls the harvest time by heating or other processing based on the effective integrated temperature. However, this method is effective only in the range where the developmental rate and temperature have a linear relationship to the effective integrated tempera-

ture, which limits the temperature that can be applied. In addition, the method is based on experience and is not based on scientific evidence. Thus, harvest time control at the present stage is still in the trial-and-error stage in experiments and production sites, and accurate harvest time control has not been achieved.

The purpose of this study is to predict the amount of growth using sensor data as a basic technology to control the harvest time more accurately. Specifically, we will predict the amount of growth by a machine learning technique from the measurement data obtained by various agricultural sensors installed in the farm. By predicting the growth of plant stems, it is thought to be possible to find factors in the environment such as temperature and humidity that are necessary for harvest time control. However, for highly accurate prediction, the training data must contain environmental conditions and patterns over the past few days that promote plant growth.

Therefore, we propose a feature generation method using agricultural sensor measurements over the past few days for predicting the amount of plant stem growth. Since plant growth does not depend on the response of the plant on only one day, but depends on the changes in the environment over the past few days, the proposed method designs features that explain the plant growth using the sensor measurements over the past few days based on the predicted date.

The structure of this paper is as follows. In Chapter 2, we describe the position of this research from related studies. In Chapter 3, the proposed method is described. In Chapter 4, the proposed method is evaluated and the results and discussion are presented. Chapter 5 summarizes this paper.

## 2 RELATED WORK

### 2.1 The Current Method for Controlling the Harvest Time

The current method for controlling the harvest time is mainly the integrated temperature control [2] using the integrated temperature method [1]. The integrated temperature method is one of the methods to predict the harvest time, and the integrated temperature control is to control the harvest time by heating or other process based on the integrated temperature calculated by the integrated temperature method. In this method, when the temperature is in the range suitable for integrated temperature method, the growth and development rate are considered to have a linear relationship with temperature, and the growth rate is considered to increase as the temperature increases. Therefore, when the temperature is not in



the suitable range, for example, in case of extremely is extreme such as on a very hot day, the prediction will be inaccurate. As for the temperature range where a linear relationship is assumed and the period during which temperature control is valid, there are currently no clear values or standards for determining these, and the only way to determine them is through trial and error in experiments in production sites.

## 2.2 Feature Extraction for Agriculture

There is a study on the process of generating yield prediction models [3]. In this study, we use a python library called *tsfresh* [4] to generate features for yield prediction. The library *tsfresh* generates various features from time series data. Although a large number of time series features are generated, the feature generation method of *tsfresh* is not based on plant physiology and is not suitable for harvest time control or growth estimation. Using the time series measurements, they generate feature values are the integrated values such as mean value, maximum value, minimum value, and standard deviation that can represent global changes. However, it is considered that the generation of appropriate feature values based on plant physiology rather than mere statistics is considered necessary for the prediction of growth.

## 3 PROPOSED METHOD

### 3.1 Overview

In the proposed method, features are created from measurements taken over the past few days in order to predict the amount of plant growth.

The purpose of this study is to predict the amount of plant growth (increment in stem height). In this study, we used sensor measurements from the past few days to design features that may be related to the amount of growth. The input sensor data is assumed to be the values of environmental sensors used in general agricultural IoT. Specifically, temperature, humidity, saturation difference, carbon dioxide concentration, and solar radiation content are used. In addition to sensor data, meteorological data released by the Japan Meteorological Agency will be used. From the meteorological data, information such as daily sunshine hours and weather conditions can be obtained. The design of features in this study can be divided into two main steps. First, as Step 1, we generate features for each day using sensor data of one day. Next, as Step 2, we generate features for a time period, which are generated from the one-day features of the past  $k$  days of the predicted date.

In this way, if there are  $n$  days of observation in the data,  $n$  values can be generated for each of the period features. One feature dataset is basically generated not per plant, but per field because environmental sensor values are common to all plants in the same field. Thus even if the dataset contains values for  $n$  days, the data size is generally small for machine learning. For this reason, we apply the machine learning method called Random Forests, which can be applied to relatively small data size.

### 3.2 Step 1: Features of a Day

First, we describe the method of generating feature values for a day. Basically, we do not generate feature values by mixing different types of measurements. Instead, we generate a feature value from a time series single measurement item. For example, we generate such as generating the maximum temperature from all the measurements of temperature in a day. As mentioned earlier, we use five items of measurements: temperature, humidity, saturation difference, carbon dioxide concentration, and solar radiation as well as the weather announced by the Japan Meteorological Agency. The weather is classified into three types: sunny, rainy, and cloudy, based on the weather announced by the Japan Meteorological Agency. The saturation difference is the value that indicates how much additional moisture can be contained in air, and is calculated from the temperature and humidity.

Weather can be used as a feature value for a day. The other values are time-series quantities, which are measured at regular time intervals. These time-series measured quantities are converted into daily feature values by the following three methods.

- (i) Levels of maximum, minimum, and average values.
- (ii) Total time for which the value takes within a specific range.
- (iii) Others.

(i) We divide the range of each measured value into multiple segments, and expresses the level as one of the segments in which the maximum, minimum, and average values are of a day included. First, let the sensor type be  $s \in S$ , and write  $v_{max}^{d,s}$ ,  $v_{min}^{d,s}$ , and  $v_{avg}^{d,s}$  for the maximum, minimum, and average values of sensor  $s$  on date  $d \in D$ , respectively. Next, we divide the range of the sensor  $s$  measurements into  $|L^s|$  segments where  $L^s$  is the set of segments and  $|L^s|$  is the number of elements. Boundary values  $b_0, b_1, \dots, b_{|L^s|}$  are defined, and when a certain measured value  $v^s$  of sensor  $s$  satisfies  $b_{l-1} < v^s < b_l$  ( $l = 1, 2, \dots, |L^s|$ ),  $v^s$  belongs to segment  $l$ . Considering this, the segment to which the maximum value  $v_{max}^{d,s}$ , minimum value  $v_{min}^{d,s}$ , and average value  $v_{avg}^{d,s}$  of each sensor  $s$  belong is considered to be a feature of a day, respectively.

(ii) We use the same segments as (i), and the total time that the daily measurements belong to each segment is used as the feature value. Let the measurement time interval of each sensor  $s$  be  $T^s$  seconds. If a measured value  $v_t^s$  at a certain time  $t$  belongs to the segment  $l$ , then  $\frac{86400}{T^s}$  seconds of the day belong to the segment  $l$ . In other words, from the number of measurements belonging to  $l$ , we can calculate the total time accumulated in  $l$ . For all combinations of each sensor  $s$  and each category  $l \in L^s$  of its values, the total time that the measurements of a day are included in the segment is the feature of the day.

(iii) As other feature values, we define values using weather data. The weather of each day (sunny, cloudy, rainy) is obtained from the Japan Meteorological Agency's web page [6] and used as a feature value. The temperature at the time of

sunrise is also obtained and used as a feature, and the category to which the temperature at sunrise belongs is also used as a feature of the day.

### 3.3 Step 2: Feature of the Period

We generate features for a period by aggregating the features of a single day in the past  $k$  days. The features of a day can be classified into numerical values and categorical values. For example, the temperature at sunrise time is a numerical value. On the other hand, the weather of each day is a categorical value. The level of the maximum value of daily measurements is a categorical value whose magnitude can be compared, and in this research, we consider it as a categorical value and generate features.

For numerical values, for a given one-day feature value, the number of days (frequency) belonging to each segment  $l \in L^s$  and the value representing continuity to be in  $l$  are defined as the feature value. Specifically, if the value of the one-day feature value of  $i$  ( $1 \geq i \geq k$ ) days ago is  $v^i$  for the feature value of one day, the corresponding period feature values are generated by each of the processes listed below.

- (1) Maximum number of consecutive days in which  $v^i$  is included in segment  $l$ .
- (2) Maximum number of consecutive days in which  $v^i$  is included in segment  $l$ .
- (3) Total time that  $v^i$  is included in the segment  $l$ .

Furthermore, we generate period features for the past  $k$  days of the following statistics from  $v_{max}^{d,s}$ ,  $v_{min}^{d,s}$ , and  $v_{avg}^{d,s}$ .

- (4) Mean value
- (5) Maximum value
- (6) Minimum value
- (7) Variance
- (8) Standard deviation
- (9) Difference between maximum and minimum

For category values, the number of days (frequency) belonging to each category and the value representing continuity to be in each category are defined as feature values, taking plant physiology into consideration. Specifically, let  $v^i$  be the value of the daily feature value of  $i$  ( $1 \geq i \geq k$ ) days prior to the predicted date, and let  $c \in C$  be the target category, the corresponding period feature values are generated by each of the processes listed below. Note that  $C$  is a set of categories.

- (a) Number of days that  $v^i$  is in category  $c$ .
- (b) Maximum number of consecutive days that  $v^i$  is in category  $c$ .

The above features are generated by the proposed method.

## 4 EVALUATION

### 4.1 Evaluation Method

Using the features created by the proposed method and the dataset of plant stem growth of the target variable, we evaluate the accuracy of predicting plant stem growth by Random Forest. The datasets to be applied are sensor data in greenhouses growing *Lisianthus* at Horticultural Experiment Center, Wakayama Agricultural Experiment Station in FY2017 and FY2020. In addition, the weather and sunshine hours for Wakayama Prefecture, where the *Lisianthus*-growing greenhouses are located, were obtained from the Japan Meteorological Agency's web page [6]. The number of days in the past for feature generation was set to 10, 7, and 3 days. In addition to the features generated by the proposed method, the number of elapsed days, which is the number of days since the start of measurement, was used as a feature. The prediction accuracy was evaluated by the leave-one-out cross-validation for the data set. The objective variable is the average of the growth for the three days leading up to the predicted date. The mean squared error was used as an evaluation index.

### 4.2 Dataset

In this section, we describe the actual data to be used and the data set created by the proposed method. We used sensor data from Horticultural Experiment Center, Wakayama Agricultural Experiment Station, image data of *Lisianthus*, and data of sunshine duration and weather from the website of the Japan Meteorological Agency. First, we used sensor data of temperature, humidity, and saturation deficit from the end of September to the end of November in 2017, and temperature, humidity, saturation deficit,  $CO_2$  concentration, and solar radiation from the end of September to the end of November in 2020. The sensor data were measured in 10-minute cycles. From this sensor data, we generate features in the proposed method.

Among the temperatures obtained from the sensor data, the temperature at sunrise time was collected, and the feature creation of the temperature at sunrise time was conducted. As a categorical data, weather was used to create the feature values. The weather was specifically divided into morning and afternoon weather, and the three categories were sunny, rainy, and cloudy.

For the objective variable "the amount of *Lisianthus* stem growth," measurements of *Lisianthus* stems were taken from the image data at a predetermined time 8:00 AM each day, and the average of the three individual stem lengths was used as the stem length of the green house on a given day. The difference between the length of the *Lisianthus* stems on a certain day and that of the previous day was used as the daily growth amount. We thought that the daily growth of *Lisianthus* stems would have errors in measurement, so we used the average of the growth of *Lisianthus* stems on one day and the growth of stems on the previous day and the next day as the objective variable. The total number of features created by the proposed method, including the number of days elapsed from the begging date of the measurement of *lisianthus* stems were in

a 1177 features. Also, the number of samples, i.e., is 47 days for the data of 2017 and 52 days for the data of 2020.

### 4.3 Evaluation Results of 2017

In this section, we describe the results of prediction by Random Forest using the data set as FY2017. The mean squared error (MSE) for each day during the whole experiment is shown in Fig. 1. For the prediction results, the average value of the MSE for all 47 days was 0.298. The overall MSE was relatively small because the mean value of growth was 1.027 and the mean value of predicted growth was 0.983. However, the error was significantly larger on the first and second days during the whole experiment. As shown in Fig. 1, there were a very few days with a large amount of growth, which were the first and second days, so the learning process did not work well. As results showed large errors in some places, it is said that the proposed method has room to improve.

### 4.4 Evaluation Results of 2020

In this section, we describe the results of prediction by Random Forest using the data set as FY2020 data. The MSE for each day during the whole experiment is shown in Fig. 2. For the prediction results, the average value of the MSE for all 52 days was 0.298. The mean value of growth is 1.745 and the mean of the predicted values is 1.761, which is a slightly smaller error than the predicted results of the FY2017 data. In addition, the maximum error was 1.426, so there was no extremely large error. However, there were some areas where the MSE was locally large, although it is not as large as the results of FY2017. One possible reason is lack of features to follow the rapid fluctuation of the growth amount.

### 4.5 Discussion

Both the results for FY2017 and FY2020 showed that the overall mean squared error was small, but the error in the region where frequent fluctuation appears was relatively large. This suggests that we have found features that can explain rough trend in growth volume, but have not generated features that can explain rapid fluctuations. Finding features to catch up with rapid fluctuations will be one of the challenging future task.

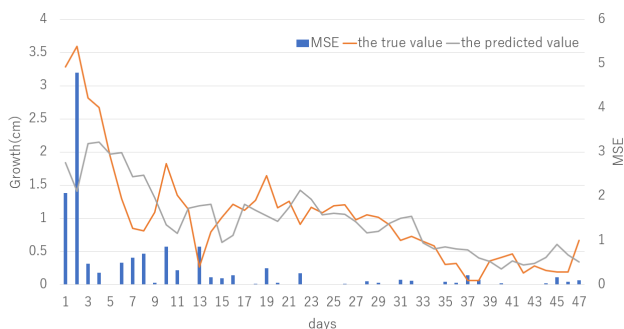


Figure 1: MSE and true and predicted values of growth in 2017

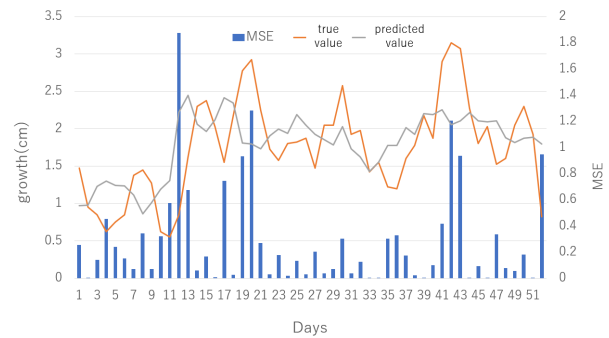


Figure 2: MSE and true and predicted values of growth in 2020

## 5 CONCLUSION

In this paper, we proposed, a feature generation method that generate a set of features from a row data set of the past several days for plant growth prediction. We applied them to a Random Forest to predict the amount of daily plant growth. By predicting the timing of the transition from nutritional growth stage to reproductive growth stage by predicting the amount of growth, we would more likely to be possible to control the harvest timing. In the evaluation, we applied the sensor data to the proposed method to generate features, and measured the prediction accuracy of the amount of *Lisianthus* stem growth with Random Forest. The prediction results of the 2017 data and the 2020 data showed small errors in the overall trend, but there was room for improvement in the prediction accuracy because the method was not able to capture the increasing or decreasing trend of the amount of growth.

## REFERENCES

- [1] E. Morie, "Effective Heat Unit Summation and Base Temperature on the Development of Rice Plant : I. A method determining base temperature and its application to the vegetative development of rice plant", Vol. 59, No. 2, pp. 225–232 (1990).
- [2] E. Heuvelink, T. Kierkels, "Plant Physiology for Environmental Control", Agriculture Japan Society of Agriculture, (2017).
- [3] Y. Todate, M. Oba, M. Takamori, "Development of Yield Prediction Model Generation Process for Fruit Vegetables in Plant Factories", IPSJ SIG Technical Report, Vol.2021-IS-155, No.7, pp. 1–8 (2021).
- [4] M. Christ, N. Braun, J. Neuffer, and A.W. Kempa-Liehr, Time Series Feature Extraction on basis of Scalable Hypothesis tests (tsfresh – A Python package). Neurocomputing, Vol. 307, pp. 72–77(2018).
- [5] G. Nakanishi, H. Mineno, "Investigation of harvest quality prediction method", The 81st National Convention of IPSJ, pp. 155–156 (2019).
- [6] Japan Meteorological Agency: Search for historical weather data, <http://www.data.jma.go.jp/obd/stats/etrn/index.php>, (2020-10-30).

# Remote monitoring system for mushrooms using LoRa communication

Hikaru Yabe\* and Mikiko Sode Tanaka\*\*

\* Kanazawa Institute of Technology, Japan

\*\* International College of Technology, Japan

\*b1800744@planet.kanazawa-it.ac.jp

\*\*sode@neptune.kanazawa-ac.jp

**Abstract** – “Nototemari” is a representative agricultural product of Ishikawa Prefecture. It is cultivated in a greenhouse in Noto Satoyama. In this paper, we report the results of a study on the construction of a remote monitoring system for Nototemari mushroom cultivation. The Nototemari mushroom cultivation greenhouse is far away from the office where the cultivation is managed, so it is costly to visit the greenhouse to monitor the growth. In addition, many producers are elderly. Therefore, we thought that we should improve the efficiency of cultivation work at low cost by constructing a LoRa network and periodically sending image data of the growth status to the management office. However, due to the slow transmission speed of LoRa communication, it took a long time to send the image data as it is. Therefore, we have built a mechanism to transmit images in a practical time by recognizing important monitoring points using the pattern recognition based on machine learning, not degrading the image quality of that part, and transmitting the image with degrading or trimming the image quality of other parts. In this paper, we show the experimental results of this remote monitoring system and show its effectiveness.

**Keywords:** LPWA, LoRa, Production control, Agriculture, Image data

## 1 INTRODUCTION

In Ishikawa Prefecture, producers, agricultural organizations, markets, prefectures, etc. have come together to brand the shiitake mushrooms as Nototemari, which is suitable for the climate of Noto Satoyama, and is located in the Oku Noto area. However, since the Noto Satoyama, especially the greenhouse that grows Nototemari, is located in the mountains of Noto, it is a blank area of radio waves and no communication environment [1]. Figure 1 shows the Noto Satoyama image illustration. In addition, it is difficult to get the power supply. Therefore, it is not possible to use a remote monitoring system that uses cameras and sensors, and all cultivation is carried out manually. Many producers are elderly, and it is desired to reduce the man-hours for visiting and monitoring to a Nototemari mushroom greenhouse.

In recent years, the use of LPWA, which does not require communication charges, has been progressing. No application is required as it uses a license-free frequency band. In particular, the private LoRa allows you to install the base station yourself, enabling flexible system design. In addition, 250mWLoRa has a long transmission distance and

is often used in mountains [2]. The problem is that it has a narrow bandwidth and very little data can be sent. It is possible to send sensor data such as temperature and humidity in a short time, but it is very difficult to send images and videos. Therefore, research is being conducted to send images with LoRa[3, 4, 5, 6].

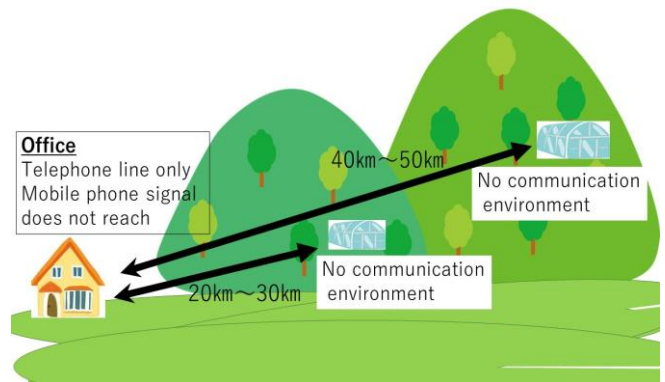


Figure 1: Communication environment of Noto Satoyama

In LoRa technology, total transmission time is limited from the viewpoint of the occupancy prevention and affects to the entire transmission time for large data like a image file. Therefore, it is said that it is difficult to send image and video data. In addition, generally, LoRa platform assumed to be inappropriate to dispatch high bit rate data, such as image or voice, due to its narrow bandwidth accessible for physical layer modulation. Bandwidth (bw) and spread rate (sf) can be changed with Lora technology. The narrower the bandwidth and the larger the spread rate, the longer the transmission time and the higher the power consumption. In other words, it takes time to transmit data to fly far.

The first attempt to transfer image data over a LoRa network was presented in [4] by C. Pham in 2016, a year after the introduction of the LoRa framework. C. Pham proposed a low cost and low power visual supervision platform based on image compression and a change detection technique. The image compression they implemented a packet loss-tolerant image compression technique that can run on very limited memory platforms. The experimental results from the tests showed that an image of about 2.4 KBytes could be transmitted up to 1.8 km. Although in this work, it is difficult to be used for external surroundings with continuously variation of brightness.

Chen et. al. [5] proposed a new trustworthy communication protocol called MPLR for image

dispatching in LoRa. It facilitates image monitoring in an agricultural IoT platform. The MPLR protocol groups information packet transmissions. By returning one ACK for each group, the time required for ACK is reduced. The test results showed that MPLR protocol during the image dispatching procedure decrease the time by 24%. This method is effective when there are few data transmission errors. Therefore, it is not suitable for large size data.

Ji et al [6], proposed a method in which doesn't transmit full images due to reduced data rate and bandwidth. Full images are not needed to be dispatched, and by this way bandwidth usage on LoRa can be reduced. The scheme took advantage of the little daily change of various growth parts of crops. And especially in farming by suggesting a new monitoring plan that splits every image into tiny grid patches. But they observed that most points in an image should be static where significant changes are very rare. Therefore, in reality, the change was so large that it could not be used.

In order to solve the problem of the conventional method, we propose a new method to divide the photo into the part that the farmer wants to see and the part that is not so important, and transfer it with reduced image accuracy of the part that is not important. The proposal system uses AI image recognition to discover important points. This makes it possible to increase the rate of image compression. By reducing the image size, the load and time required for transmission can be reduced, and a practical system can be constructed.

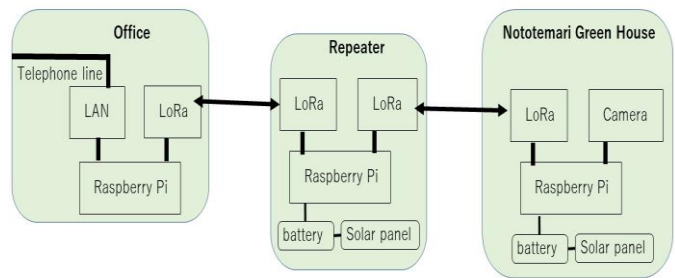
In this paper, we explain the communication infrastructure using LoRa that is the basis of the remote monitoring system. Also we explain a protocol and a frame data set for the system. In addition, we explain the result of the conduct a demonstration experiment.

## 2 SYSTEM CONFIGURATION

The system configuration of the remote monitoring system for Nototemari will be described. In the greenhouse, in addition to temperature and humidity, images are taken once a day and sent to the office by LoRa. Due to the distance between the greenhouse and the office, there are several repeaters, and the data is sent to the office via the repeaters. Figure 2 shows the overall image of the remote monitoring system for Nototemari.

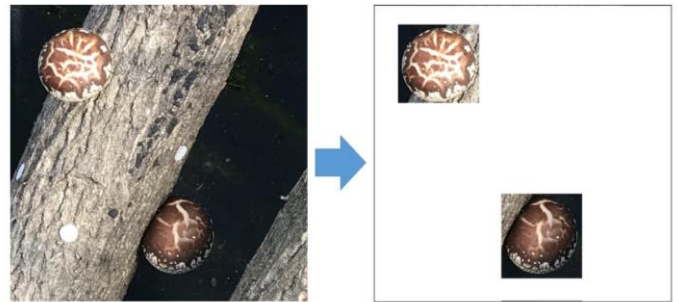
Since the office is located in a mountainous area, only telephone lines are often available, and the system is such that data is uploaded to the cloud using telephone lines. Since the greenhouse has no power source, it secures electricity by solar power generation. Therefore, the power is turned off except for acquiring data once a day and transmitting it. The repeater holds two LoRa modules. One is for reception and the other is for transmission. With these two LoRa modules, data can be sent to the next node without stagnation. Depending on the distance between the office and the greenhouse, multiple repeaters may be placed.

Temperature and humidity are important factors for the growth of Nototemari [7]. In addition, there are many items that should be visually confirmed, such as how the umbrella is wrapped, and images are important for growth



management [8]. For these reasons, a temperature / humidity

Figure 2: System configuration of remote monitoring system for "Nototemari"



sensor and cameras were installed in the greenhouse, and these data was transmitted to the office using the LoRa

Figure 3: A example of image data compression  
(Right: Before compression, Left: After compression)

network so that it could be confirmed at the office. Since LoRa has a narrow band, it is difficult to send image data as it is. Therefore, we decided to use the pattern recognition based on machine learning to extract only the necessary parts and send them.

We succeeded in compressing the amount of data while maintaining the image quality required for growth management. Figure 3 shows an example of image data compression. The left is the data before compression, and the right is the data after compression. In this example, compression is used to erase all data or to reduce image quality except the important Nototemari. The compression level can be specified. Items that perform AI recognition can also be controlled. It's important to be able to control, as each growth process has different items to check.

In order to recognize the important part, acquire the training image data in advance, train these, and create a library. This library is used to recognize important parts, maintain the image quality of those parts, and compress the image size by transmitting the other parts with reduced image quality. Next, in order to perform the compression in H.265, change to the video data, performs compression in H.265 [9], it transmit and receive data at LoRa. If the image data size is large, it will take time to send. In addition, the possibility of transmission errors increases, which is not desirable for the system. Therefore, we decided to send the high accuracy only the parts that the producer must confirm with high accuracy. This makes it possible to send in a few minutes, which is practical.



### 3 COMMUNICATION PROTOCOL AND FRAME STRUCTURE

The protocol for transmitting image data and the data frame will be described. Figure 4 shows an image data transmission frame structure. With LoRa, a maximum of 255 Bytes can be transmitted with one transmission. The first 24 Bytes is the module-specific usage area we are using. Next, the payload comes, and the image data is put in this area and transmitted. The communication protocol used in the image file transmission is the stop-and-wait scheme [5]. In the stop-and-wait scheme, sender send the data, then the sender waits for an acknowledgement per packet to ensure that the data arrives correctly. If sender get the ACK, sender send the next data. If sender don't get the ACK, sender send the same data.

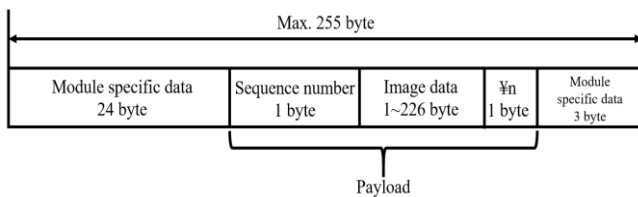


Figure 4: Frame Structure for image data

The problem in implementing the stop-and-wait scheme is that it cannot determine the duplication of packets. If an Ack sent from the receiving side to the sending side is lost or corrupted, the sending side times out and retransmits the frame. In this case, the receiving side will have two frames with the same content, and the data will not be consistent. To solve this problem, a sequence number is defined in the first Byte of the frame, and 0 and 1 are entered alternately for each transmission. This allows the receiving side to detect duplicate frames by checking if the sequence number is entered alternately.

When there is no more data to send, the sender sends an exit character. The receiver receive the exit character, completes the reception and sends an Ack to the sender. This completes the sending and receiving process. After that, the receiver continues to wait for data to be received.

### 4 EXPERIMENTAL RESULTS

We will explain the experiment of image degeneracy using he pattern recognition based on machine learning and the experiment of transmitting the data after degeneracy with LoRa. The result of degeneracy using he pattern recognition based on machine learning is shown for the photograph taken at Nototemari greenhouse. The photo on the left of Figure 5 is taken with a camera. The photo on the right of Figure 5 is after reduction by the pattern recognition based on machine learning. In this example, the part other than the part recognized by the pattern recognition based on machine learning is painted white. File size of the original image is 4,164,000 Bytes. And file size of extraction by the pattern recognition based on machine learning is 352,907 Bytes. The size become about 1/10. After H. 265

compression, the file size is 69,955Bytes. It has been degenerated to one-fifth. The size of each file is summarized in Table 1. As can be seen from the figures, compression does not have a significant impact on the clarity of the images.

Compression using the pattern recognition based on machine learning is different from the conventional method of reducing the image quality of the entire image, and it is possible to acquire high-quality data for the part you want to see. In addition, farm workers handle the rotation of logs of Nototemari. In the proposed method, there is no problem even if the position of the Nototemari changes in the camera image. Also, if you register a disease etc. in the AI library, the data will be sent with a clear image of that part, so you will not overlook it.

H.265 has the disadvantage of enormous calculation costs for encoding time and compression ratio. In fact, the compression process to H.265 takes a few minutes on a Raspberry Pi. However, considering the transmission time in LoRa, we thought that data should be compressed by edge computing even if it takes time.

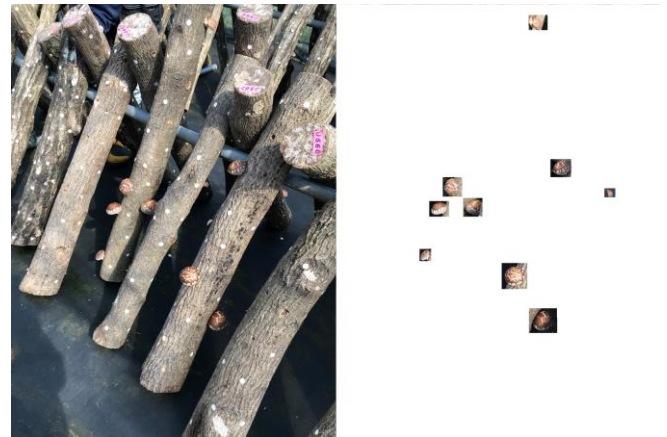


Figure 5: Result of the data compression  
(Right: Original image data Left: Extracted image data by the pattern recognition based on machine learning)

We tested how long it takes to transmit a compressed image. Figure 6 shows the experimental environment. The experiment was conducted indoors. We prepared the equipment to be installed in the greenhouse and the gateway, and communicated between them. Figure 7 is an image acquired by the gateway. You can see that there is no image deterioration compared to the transmitted data. In addition, since the parts that users want to observe with the image quality taken by the camera are retained, it can be said that it is sufficiently practical for growth management of Nototemari. 4 cameras was installed in the greenhouse. The time required to transfer the data of one image was about 10 minutes. Thus all the data is sent, it will be within 1 hour, so the system can send the data once a day. Then, it can withstand the operation with the solar panel and battery.

Table 1: Image Related Information

	Data size
Original image	4,164,000 Bytes
Extraction by AI	352,907 Bytes
After compression with H.265	69,955 Bytes

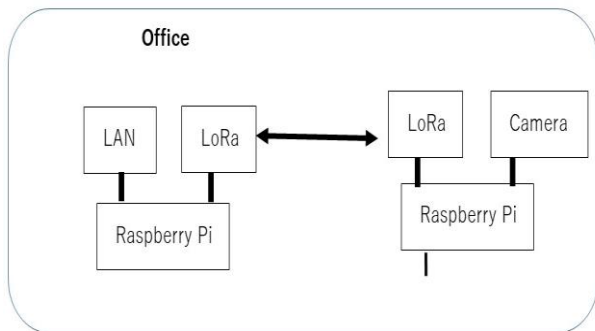


Figure 6: LoRa transmission test environment



Figure 7: Received image data

## 5 CONCLUSION

Nototemari is a representative agricultural product of Ishikawa Prefecture. It is cultivated in a greenhouse in Noto's Satoyama. Noto Satoyama has no communication environment at all. In this paper, we reported the results of a study on the construction of a remote monitoring system for Nototemari cultivation. The Nototemari cultivation greenhouse is far away from the office where the cultivation is managed, so it is costly to visit the greenhouse to monitor the growth. In addition, many producers are elderly. Therefore, we thought that we could improve the efficiency of cultivation work at low cost by constructing a LoRa network and periodically sending image data of the growth status to the management office. However, due to the slow transmission speed of LoRa communication, it took a long time to send the image data as it was. Therefore, we have built a mechanism to transmit images in a practical time by

recognizing a Nototemari using the pattern recognition based on machine learning, not degrading the image quality of that part, and transmitting the image with degrading the image quality of other parts. In this paper, we show the experimental results of the propose image data sending system, and show its effectiveness. The proposed an image data sending system turned out to be sufficiently practical because one image data can be sent in a sufficient degree.

## REFERENCES

The research is supported by the Telecommunication Advancement Foundation.

## REFERENCES

- [1] Adhere to radio wave countermeasures on the Noto Peninsula! What is a device that suits the mountains, sea, sightseeing spots, and the natural environment?, <https://time-space.kddi.com/au-kddi/20200305/2854>, Access 2021.5.7.
- [2] Junji Moribe, Akifumi Fujimoto and Yoshiaki Tokita, "Development of a data notification system using GEO-WAVE," IEICE Communications Express, Vol.8, No.12, 536–541.
- [3] Anestis Staikopoulos, Venetis Kanakaris, George A. Papakostas, "Image Transmission via LoRa Networks – A Survey," 2020 IEEE 5th International Conference on Image, Vision and Computing (ICIVC), Beijing, China.
- [4] C. Pham, "Low-cost, low-power and long-range image sensor for visual surveillance," AMC SmartObject '16, pp. 35-40, 2016.
- [5] Chen, T., Eager, D. and Makaroff, D., "Efficient Image Transmission Using LoRa Technology In Agricultural Monitoring IoT Systems," In 2019 International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing(CPSCom) and IEEE Smart Data (SmartData) (pp. 937-944). IEEE, 2019, July.
- [6] Ji, M., Yoon, J., Choo, J., Jang, M. and Smith, A., "LoRa-based Visual Monitoring Scheme for Agriculture IoT," In 2019 IEEE Sensors Applications Symposium (SAS) (pp. 1-6). IEEE, 2019, March.
- [7] Yashirna T, Kodani J, Kado M, "Influence of different cultivation environment in vinyl house to fruit body development of the large shi i take mushroom," <https://agriknowledge.affrc.go.jp/RN/2010922309.pdf>
- [8] Notomari cultivation guide, <https://www.pref.ishikawa.lg.jp/ringyo/publish/documents/nototemari2018.pdf>, access 2021.5.7.
- [9] <http://ffmpeg.org/>, access 2021.7.24.
- [10] [http://www.rflink.co.jp/pdf/RM-92A-92C/SimpleMACstd92A-92C\\_instruction%20manual-rev2.9.15.pdf](http://www.rflink.co.jp/pdf/RM-92A-92C/SimpleMACstd92A-92C_instruction%20manual-rev2.9.15.pdf), access 2021.7.24.



Keynote Speech:

Dr. Akio Yamada

( Senior Vice President and  
Head of NEC Laboratories )



## Digital innovation towards sustainable society

Akio Yamada \*

\*NEC Corporation, Japan

**Abstract** - Several advanced Information and Communication Technology (ICT) and its latest applications are presented. ICT has a big potential to innovate human society by replacing or combining with various conventional technologies. In the past, the mainstream of digitalization was aimed at improving business productivity and efficiency. However, it has shifted to create greater social value, such as changing the lifestyles of consumers and solving social issues. This is a common trend for major ICT vendors and a phrase, “innovating society to be sustainable” is widely spread. Sustainability is one of the most frequently used keyword in these days. Various agendas are studied, for example, energy innovation, fair access to education service, food waste reduction, and so on. Digital transformation of society/industry using advanced ICT is considered as the most promising approach to solve the problems raised in the agendas. In this presentation, various advanced technologies are introduced, including how they can contribute for making sustainable world. First, to be healthy and comfort daily life, three challenges are discussed; biometric National ID to deliver fundamental service for everyone around the world, AI assisted drug discovery to realize personalized medicine, and novel shoes-embedded sensors to find a sign of health problems. To be safe and efficient society, three trials are demonstrated; automated operation of complex social infrastructure to optimize its usage, secure data analysis to realize both protection and utilization of private data, and prescriptive analysis and automated actuation to real-world using robots. Finally, to be environment friendly society, three deployments are introduced; new energy saving cooling system for large scale data centers, AI aided farm management to increase productivity, and city-wide environment monitoring for smart city operation.



Akio Yamada received Ph.D. in electronic and information engineering from Nagoya University and joined Central Research Laboratories of NEC Corporation in 1993. He started his research career in digital media distribution system and expand it to media content recognition, ICT system architecture, and knowledge discovery science. After having business experiences in enterprise DX market as Vice President on Technology, he is now Senior Vice President and Head of NEC Laboratories. He has also contributed to many international standards in media content distribution and processing area, known as MPEG and JPEG, for about 20 years and received a lot of awards.

IWIN2021 Keynote

# Digital innovation towards sustainable society

Sep 13, 2021

Akio Yamada

Senior Vice President (Corporate R&D), NEC Corp.

© NEC Corporation 2021

\Orchestrating a brighter world

NEC creates the social values of safety, security,  
fairness and efficiency to promote a more sustainable world  
where everyone has the chance to reach their full potential.

## Today's topics

### 1. Who is NEC? - Corporate Profile

### 2. Latest research and its social implementation

#### - For better life

Biometric National ID / AI assisted drug discovery / Shoes embedded sensor

#### - For innovative society

Automated operation of complex infrastructure / Secure data analysis system / Prescriptive analysis and actuation

#### - For better environment

Energy-saving DC cooling / AI aided farm management / City-wide environment monitoring

### 3. Summary

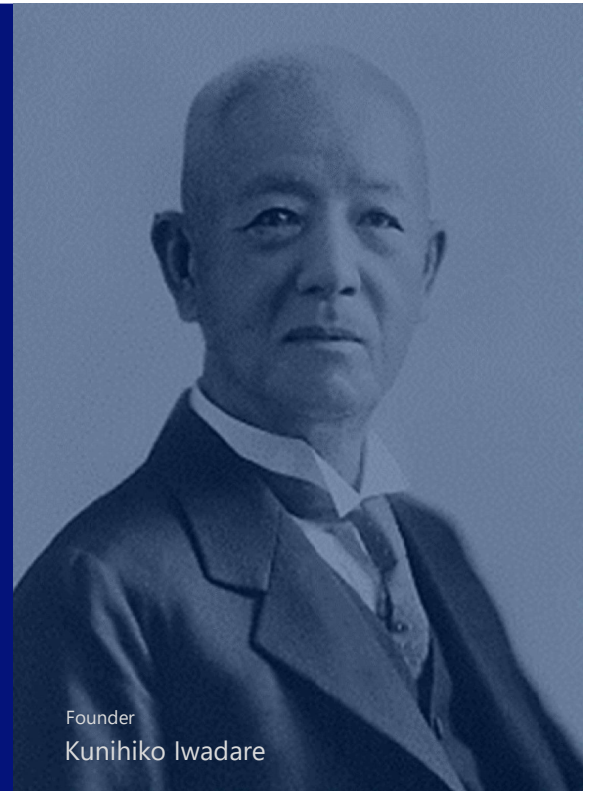
Orchestrating a brighter world **NEC**

Movie

## Our Founding Spirit

# Better Products, Better Services.

The founding spirit that has guided the NEC Group for over 120 years,  
since the company's establishment in 1899



Founder  
Kunihiro Iwadare

## NEC's Business portfolio

NEC creates the social values of safety, security, fairness and efficiency to promote a more sustainable world where everyone has the chance to reach their full potential.

**DIGITAL GOVERNMENT  
DIGITAL FINANCE**

**GLOBAL 5G**

**Strong Focus**

**Biometrics**

**AI**

**5G/TOMS**

**Cloud**

**+ M&A**

**Security**

**R&D**

**Public & Communication  
Infrastructure**

**CORE DX**

**IT SERVICES &  
PRODUCTS**

**Transformation**

# NEC 2030VISION

## Life

Bringing people together and filling each day with inspiration

## Society

Nurturing prosperous cities with inclusive and harmonious societies  
Creating sustainable societies by shaping new industries and workstyles  
Sharing hopes that transcend time, space, and generational boundaries

## Environment

Living harmoniously with the earth to secure the future

## R&D Focus

Create social values by cyber physical system using cutting-edge AI and ICT platform





# Global R&D system

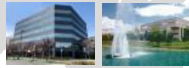
Established labs in seven locations around the world to conduct cutting-edge research and development

## NEC Labs. Europe



Create solutions and technologies through EU-PJ and high-tech social implementation

## NEC Labs. America



Develop advanced technologies by leveraging geographical advantage of state-of-the-art technology

## Israel Research Center



Rapidly create solutions in the world's leading startup country by combining external advanced technologies with NEC technologies

## NEC Labs. India



Create solutions and technologies focused on social issues in emerging countries

## NEC Labs. China



AI-related and network-related research and development

## 4 Labs. in Japan



The control tower for NEC's research and development. In addition to AI technologies (Analysis, Recognition), Security, ICT platform, focusing on advanced technologies of quantum computing and devices.

## NEC Labs. Singapore



Create solutions focusing on social issues in developed countries, based on collaboration with local governments and customers

## NEC 2030VISION

### Life

Bringing people together and filling each day with inspiration

### Society

Nurturing prosperous cities with inclusive and harmonious societies

Creating sustainable societies by shaping new industries and workstyles

Sharing hopes that transcend time, space, and generational boundaries

### Environment

Living harmoniously with the earth to secure the future

## Life

### Fair service delivery for everybody

(Biometrics National ID)



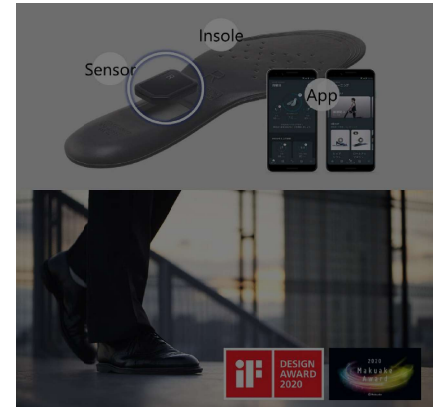
### Personalized medical care service

(AI Drug Discovery)



### Healthy life expectancy through daily monitoring

(Gait sensing and recognition)



11

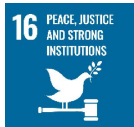
© NEC Corporation 2021

Orchestrating a brighter world **NEC**

## Maintaining peace and equity

### Identification of Newborns and Infants

- 99.7% accuracy was proven with fingerprint recognition of neonates 2 hours after birth (first in the world) in Kenya.
- In collaboration with Gavi, the Vaccine Alliance, NEC is promoting the practical application of infant fingerprint recognition aimed at the proliferation of vaccines.
- NEC aims to provide public services such as legal identification and vaccination to children in developing countries.



By 2030, provide legal identity for all, including birth registration (SDGs 16.9).



12

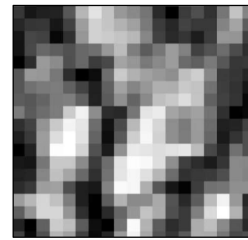
© NEC Corporation 2021

Orchestrating a brighter world **NEC**

Maintaining peace and equity

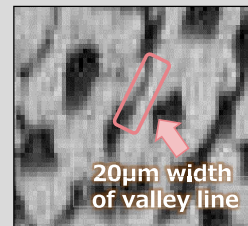
## Technical challenge

How to **reliably sample fingerprints**, which are unchanging over a lifetime, **from a neonate**



Fingerprint image captured with conventional method

**NEC optimized the FOP's fiber diameter and thickness, as well as the resolution of image sensors, and succeeded in capturing images of fingerprints of neonates within 2 hours of their birth (20- $\mu$ m valley lines) for the first time in the world.**



Fingerprint image captured using NEC's fingerprint identification technology

Won the 18th International Conference of the Biometrics Special Interest Group (BIOSIG2019) Best Paper Award

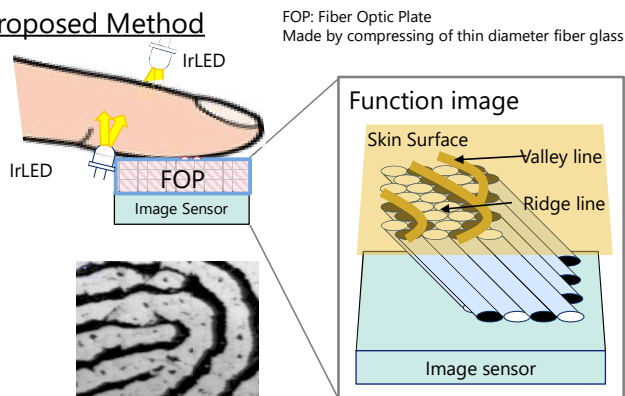


Movie

### Maintaining peace and equity

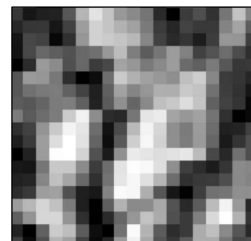
- FOP specially crafted for neonate keeps lighting intensity of ridge line as a bright area and of valley line as a dark area and leads the contrast directly to the sensing cell without interfering.

#### Proposed Method

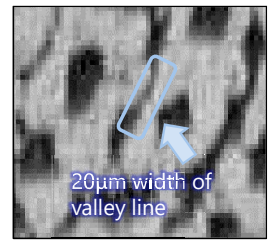


FOP leads lighting intensity directly to the sensing cell without interfering.

- The optimized combination of image sensor and FOP successfully takes the neonate fingerprint image with a line width of 20 $\mu$ m (first in the world).
- Research team built by a collaboration with Kenya Ministry of Health took more than 1,000 images.



Conventional Method



Proposed Method



This research was the best paper award at 18<sup>th</sup> International Conference of the Biometrics Special Interest Group (BIOSIG2019).

### Proving optimal care to individual patients

## AI Drug Development Business

- A clinical trial for personalized cancer vaccines has started in Europe.
- Strategic Partnership & M&A
- NEC's AI is being used for both cancer immunotherapy as well as infectious disease (e.g., COVID-19 vaccines)

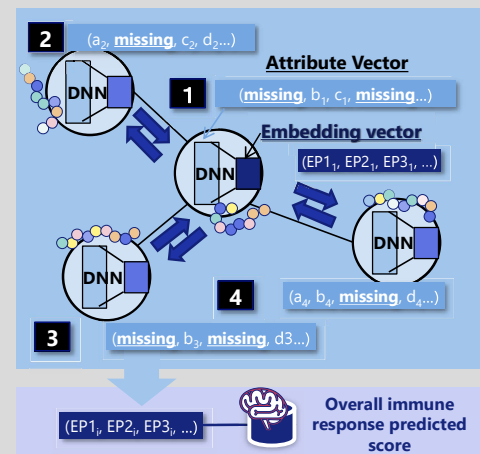


## Technical challenge

For practical use of a peptide vaccine for cancer treatment, it is necessary to **discover a peptide that activates immunity from about 500 billion amino acid sequences**. The problem is that the amount of experimental **data is not enough**.

**NEC's unique Graph based Relational Learning "Embedding Propagation" compensates missing data with high accuracy.**

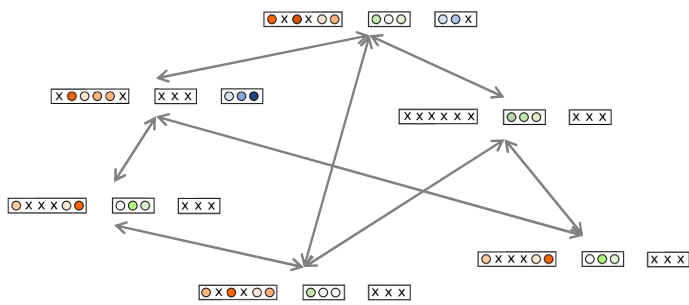
**This dramatically improves the prediction accuracy of neoantigen.**



Movie



## Proving optimal care to individual patients



- **GraphAI** creates/learns and exploits **relationships** between data points explicitly (data points “talk to each other” during learning)
- **GraphAI** can use **unlabeled data** to learn a better vector representation
- **GraphAI** provides opportunities to learn how to fill in **missing data**

- The **Embedding Propagation (EP)** algorithm is one component of our **GraphAI** portfolio
- It learns an **embedding** (a numeric vector) for each data instance based on a (possibly learned) graph structure
- The embeddings show **superior performance for downstream tasks** such as classification and regression
- EP has three key advantages compared to other methods:

1

Embeds  
graph knowledge

2

Integrates  
multi-modal data

3

Handles  
missing values

## Prolonging healthy age

## Measuring the “quality of walking” with insole sensors

- The natural everyday gait is analyzed and recorded as data without the hassle of charging, wearing, or operating the sensors. This development contributes to prolonging healthy life through advice on walking posture and health management services.
- The gait prediction model that accurately detects walking cycle achieves precision and robust gait measurement.
- Service trial is ongoing as a joint project between NEC and FiNC Technologies.



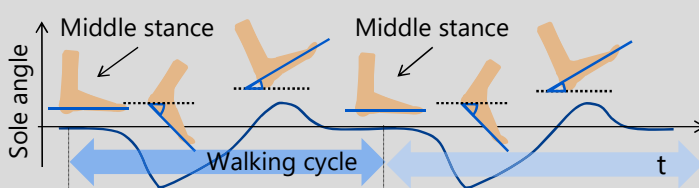
Movie

Prolonging healthy age

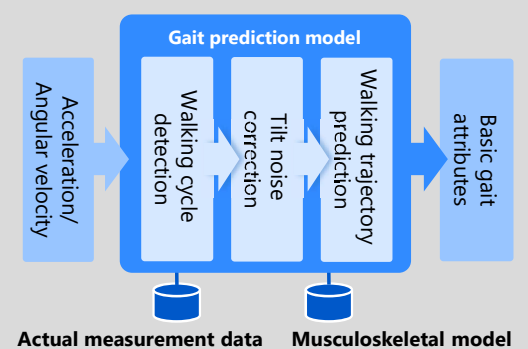
## Technical challenge

To understand health conditions, a technology is required to continuously **acquire gait data with sufficiently high precision** under various conditions of daily life **consuming limited battery**

**A gait prediction model that accurately detects walking cycle is used to correct noise in everyday measurements, achieving accurate, robust gait measurement.**



Flow for estimating gait





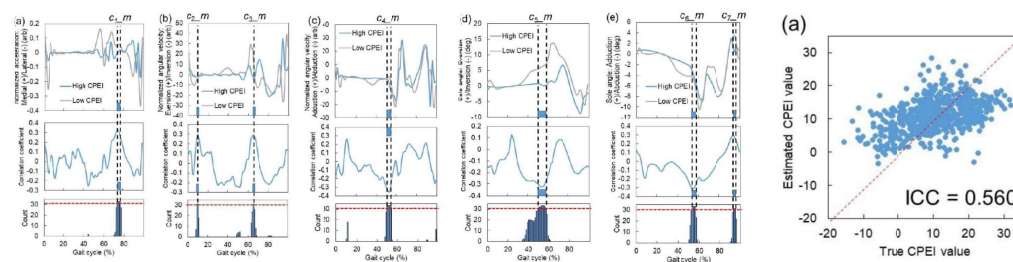
## Prolonging healthy age

"Gait analysis AI technology" that reads the health state of feet from gait data

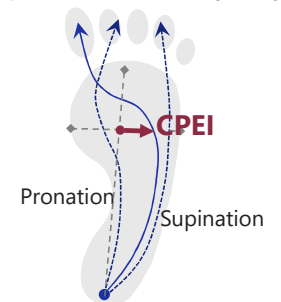
**Accurately predicting the center of pressure excursion index (CPEI), which indicate the health state of feet, from natural everyday gait data measured by insole sensors**

- Feature extraction is determined by means of a proprietary method using domain knowledge
- Excellent robustness makes possible CPEI prediction in high-noise actual-use environments
- High accuracy: Intraclass correlation coefficient (ICC) between correct values and prediction values is 0.56

Perfect shoe selection can be achieved by monitoring flatfootedness and degree of pronosupination. Health can be improved through correcting your gait.



Feet pressure-centered trajectory



Center of pressure excursion index (CPEI)

This system is not medical equipment, and thus should not be used as or for medical treatment. It is also not provided as a product as this technology is currently in the development phase.

## NEC 2030VISION

Life

Bringing people together and filling each day with inspiration

Society

Nurturing prosperous cities with inclusive and harmonious societies

Creating sustainable societies by shaping new industries and workstyles

Sharing hopes that transcend time, space, and generational boundaries

Environment

Living harmoniously with the earth to secure the future

## Society

### Efficient and stable operation of social infrastructure

(Automated optimal operation using AI)



### Strictly secure usage of privacy/confidential data

(secure computing and storage)



### Smart logistics service

(safe and high efficient robotics control)

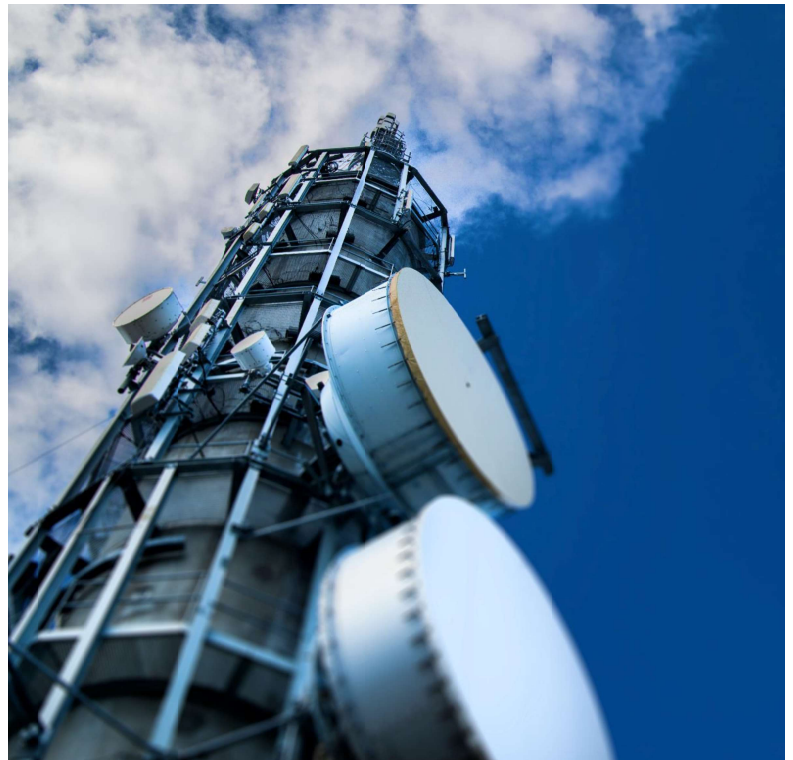


Developing communications infrastructure that meets growing demand

## Automation of establishment and operation of 5G wireless base stations

- Open architecture enables automation of base station establishment and operation (entry opportunity for new vendors).
- Cutting-edge AI is used to optimize O-RAN\*-based wireless base station operation.

\* O-RAN: Multi-vendor wireless base station specifications

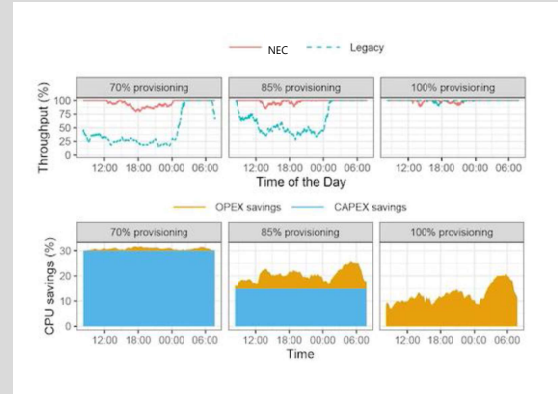


## Technical challenge

Tuning at a massive number of base stations requires man-hours and know-how. A technology that enables **automatic tuning that efficiently maximizes communications traffic capacity with low power consumption** is needed.

**Two-stage AI optimally allocates calculation resources in real-time.**

**General optimization is implemented based on data amassed over 24 hours, followed by a more detailed optimization based on subsequent changes.**



Movie



## Social stabilization by data democratization

## Gaining new findings through analyses of personal and organizational information while keeping such information private

- Integrate confidential data such as financial information and genomic analysis data in possession of individuals and organizations in order to discover new findings
- While keeping the secrecy of personal data, acquire only the results of processing such data through statistical analysis, comparison and matching, and scoring
- Keep data secure over multiple servers based on a secret sharing scheme and achieve distributed processing by multi-party computation without revealing anything but the output



29

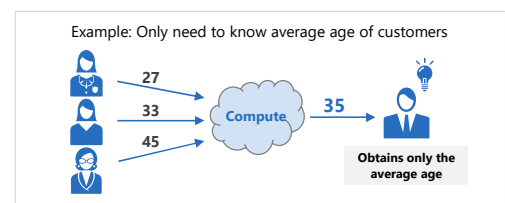
© NEC Corporation 2021

Orchestrating a brighter world **NEC**

## Social stabilization by data democratization

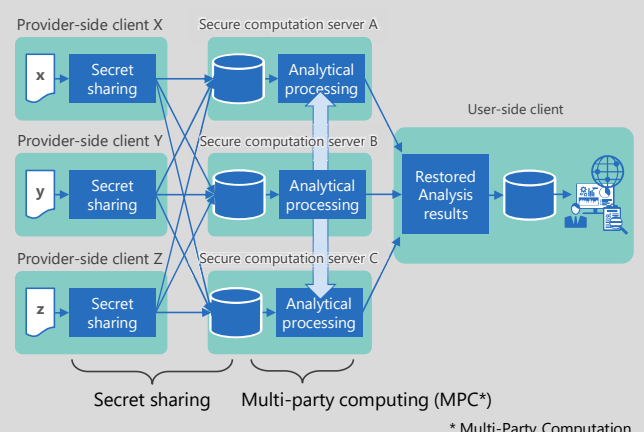
## Technical challenge

Contributing to the development of business and society through applying collective intelligence generated from **integrated processing of personal data and trade secrets while keeping them private**



## Keep data secure and preserve privacy through multi-party computation(MPC) using secret sharing scheme in data processing (e.g. statistical analysis, comparison and matching, and scoring )

During the MPC process, the input is confidential, shared data in which personal information is encrypted and split on different servers. The MPC output is also split. This scheme reduces the risk of information leakage and the cost of storage operations because each server cannot infer the original personal data.



\* Multi-Party Computation

30

© NEC Corporation 2021

Orchestrating a brighter world **NEC**

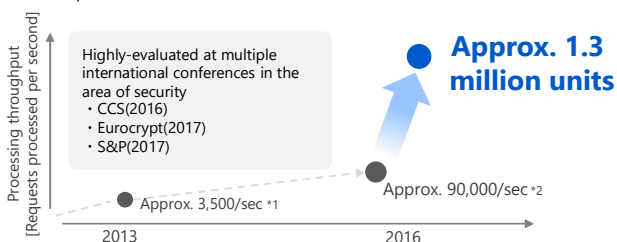
## Movie

## Social stabilization by data democratization

## Advantages of NEC's Secure Multi-party Computation(MPC)

## Achieving both high security and practicality

- Conventional secure computation had to trade off **high security performance** with **significantly increased processing time**
- NEC achieved **MPC algorithm for high-speed processing based on our original technologies**
  - AES processing throughput was conventionally maxed out around 90,000/sec, but now is **increased to around 1.3 million/sec** (ACM CCS 2016 Best Paper award)



※1 S. Laur, R. Talviste and J. Willemson. "From Oblivious AES to Efficient and Secure Database Join in the Multiparty Setting", ACNS2013.

※2 J. Randmets. Personal comm. AES performance on the new Sharemind cluster. May, 2016.

## Support tools that increase productivity

- **Data scientists not specializing in secure computation can also participate in development**
  - Advantage 1. Description of process can be **written easily**
  - Advantage 2. Processing can be **sped up easily**

## Advantage 1

- Python-like programming language is available to write and compile descriptions of secure computation processing

## Advantage 2

- Execution code is automatically optimized to speed up its execution
  - Compiler automatically detects transactions that can be executed in parallel

Source : [https://www.nec.com/en/global/rd/technologies/201805/pdf/mpc\\_introduction.pdf](https://www.nec.com/en/global/rd/technologies/201805/pdf/mpc_introduction.pdf)

Source : [https://jpn.nec.com/press/201811/20181105\\_02.html](https://jpn.nec.com/press/201811/20181105_02.html)

## Improving labor productivity

## Achieving safe, high-efficiency collaborative transport using robots (Logistics)

- Multiple transport robots autonomously predict and avoids collision in an environment where human operators are also moving around
- Transport efficiency is double that of conventional transport robots due to the avoidance of traffic congestion inside the warehouse

Real-world recognition

× Communication

× Control

× Inter-robot collaboration



33

© NEC Corporation 2021

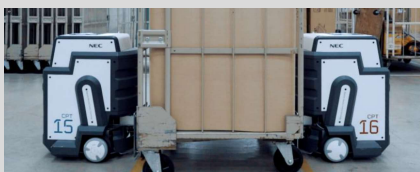
Orchestrating a brighter world **NEC**

## Improving labor productivity

## Technical challenge

To improve **transport efficiency** with minimum computational cost **while using carts already in operation**

**Two remote-controlled robots carry a load by clamping it between them**  
**Speed is automatically controlled according to the degree of collision risk**



### Higher risk

Careful transport with priority given to safety



### Lower risk

Faster transport with priority given to efficiency



34

© NEC Corporation 2021

Orchestrating a brighter world **NEC**

Movie

Improving labor productivity

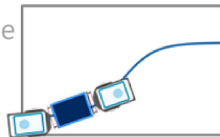
## Risk-Sensitive Stochastic Control

Break the trade-off between **safety** and **efficiency**

### Explicitly incorporate stochastic disturbance terms

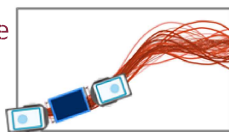
#### Conventional

Solution is unique and definite

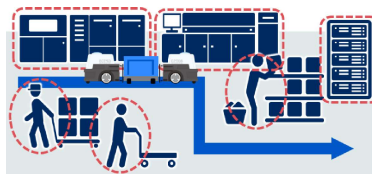


#### Proposed

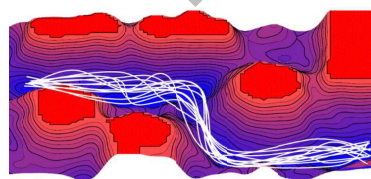
All risks could be considered



### Evaluate collision risks

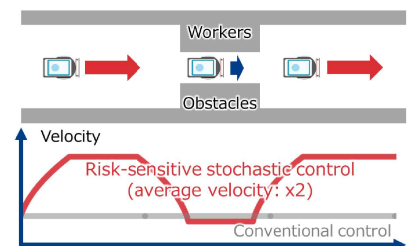


positions and motions of workers and obstacles



risk-sensitive evaluation function

### Adaptively control moving speed according to the risk function



S. Yasuda, et al., "Cooperative Transportation Robot System Using Risk-Sensitive Stochastic Control", IROS2021, to be published.



# NEC 2030VISION

## Life

Bringing people together and filling each day with inspiration

## Society

Nurturing prosperous cities with inclusive and harmonious societies

Creating sustainable societies by shaping new industries and workstyles

Sharing hopes that transcend time, space, and generational boundaries

## Environment

Living harmoniously with the earth to secure the future

37

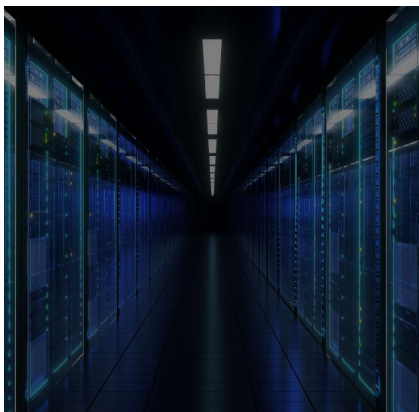
© NEC Corporation 2021

Orchestrating a brighter world **NEC**

## Environment

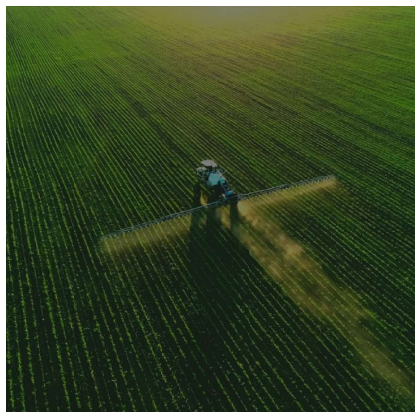
### DC operation with low power consumption

(environment friendly cooling system)



### Environment-friendly farming

(AI assistant for better productivity)



### Wide range sensing for city-scale digital twin

(multi-modal sensing using existing optical fiber)



38

© NEC Corporation 2021

Orchestrating a brighter world **NEC**

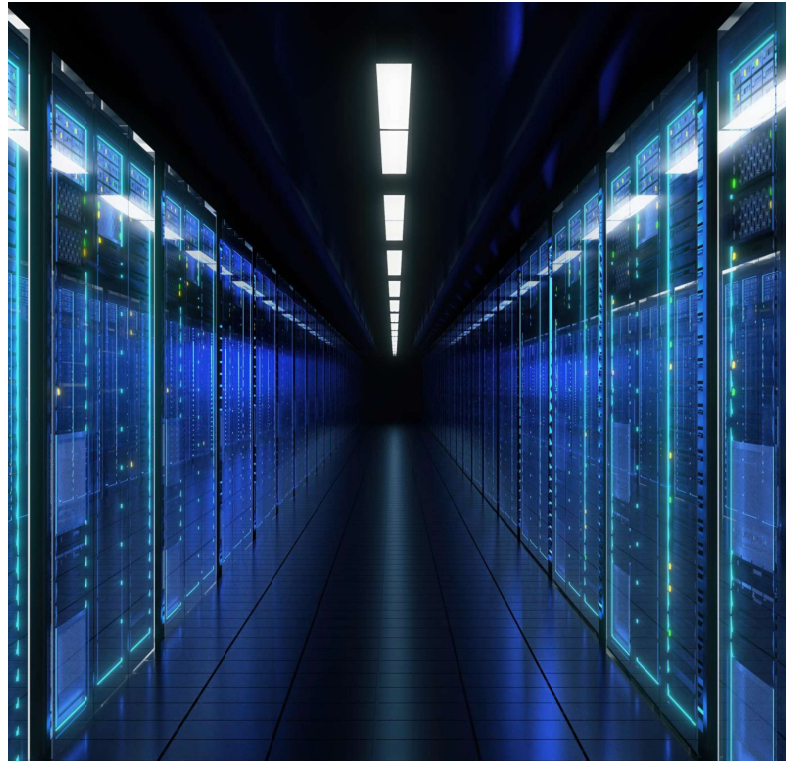
## Improving energy efficiency

## Reducing air-conditioning power in data centers to half

- Data centers are expected to grow in scale and estimated to account for 10% of global power consumption (2025)
- Power consumed by data center air-conditioning is reduced to half by heat-exchanging near the heat source using low-pressure refrigerant
- GHG reduction is also achieved by using green refrigerant



By 2030, double the global rate of improvement in energy efficiency (SDGs 7.3).



39

© NEC Corporation 2021

Orchestrating a brighter world **NEC**

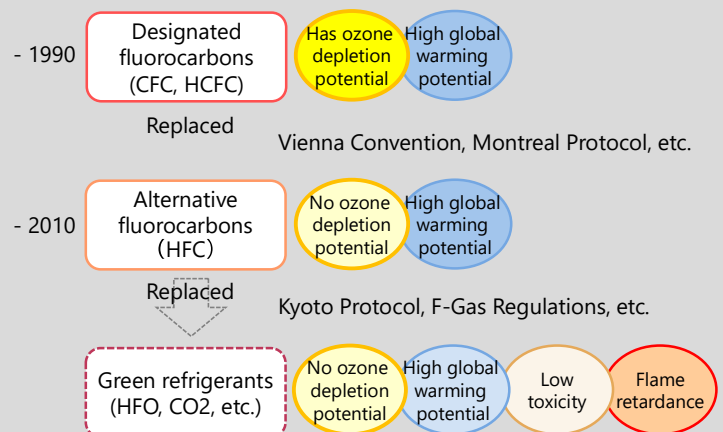
## Improving energy efficiency

## Technical challenge

CFC substitutes, the mainstream air-conditioning refrigerant, have a large environmental impact, but **high-efficiency air-conditioning** could not be achieved **with environmentally friendly green refrigerants**.

**NEC used phase change cooling technology to develop a novel air-conditioning system that uses low-pressure green refrigerant.**

**Outstanding power-saving is achieved by efficiently transferring heat from the heat source.**



40

© NEC Corporation 2021

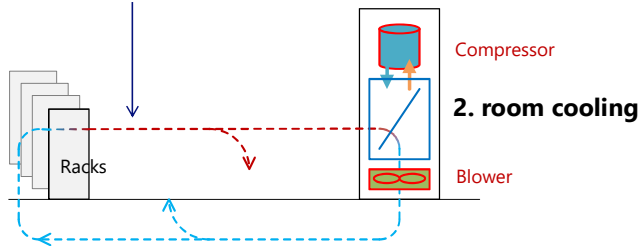
Orchestrating a brighter world **NEC**



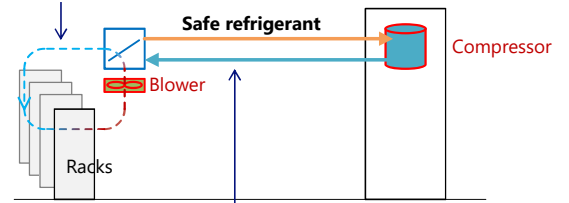
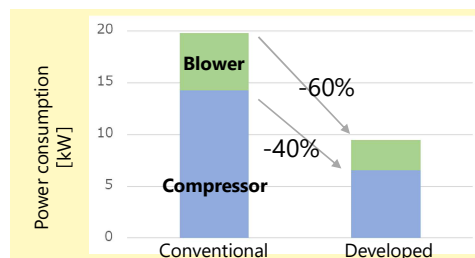
## Improving energy efficiency

## Phase Change Cooling system with Safe Non-Freon Refrigerant

## Conventional cooling system

1. **long-distance** air circulation without safe refrigerant

## Developed cooling system

1. **short-distance** air circulation -> reduce blower power2. **local cooling by heat transport**  
-> reduce compressor power

Experimental results in a real DC

## Passing on agricultural wisdom

## AI-based farming advice improves productivity and achieves green agriculture

- Amid the global drop in farming population, NEC contributes to sustainable food and agriculture by addressing issues such as technology transfers and formation of new production sites.
- Various cultivating environments are virtually reproduced, combining AI models built on codified knowledge of farming by skilled cultivators with a crop growing simulator.
- NEC provides optimal farming advice even in new places. In both northern and southern hemispheres, NEC assists with harvest of the same level as that achieved by skilled producers regardless of environmental differences.

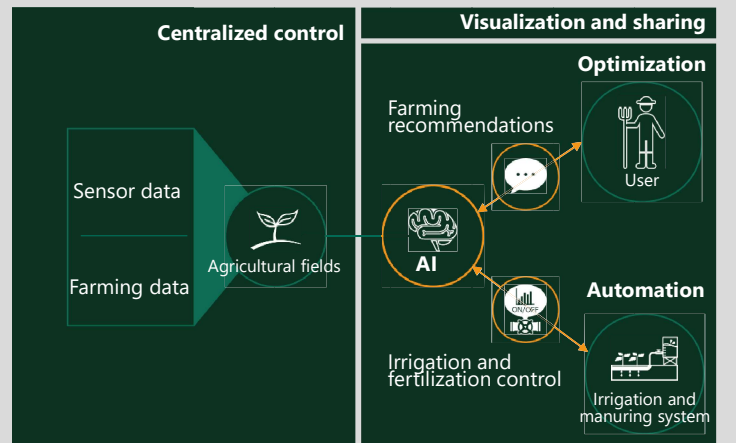


## Technical challenge

Profitable farming requires experience, time, and labor - hard to scale out.

Improvement of harvest efficiency expected over short time even with inexperienced producers.

**Crop growing model is generated by using accumulated data. AI analyzes the optimal timing and amount of irrigation and fertilization and provides recommendations.**



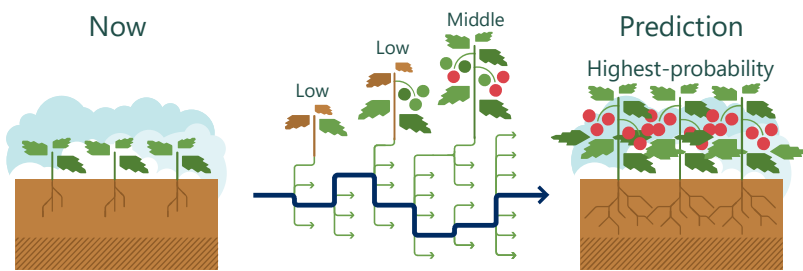
Movie

## Passing on agricultural wisdom

### Simulation prediction

#### Predicting the most probable future in simulated agricultural environment

By simulating changes in the reproduced agricultural environment, the AI singles out the most probable future out of a vast number of possibilities so that the best possible measures can be taken.



**Huge Amounts of Simulation**

Demonstration experiment conducted in Australia/Portugal

**Harvest amount increased more than 1.3 times**

Compared to the area average in Portugal

Demonstration experiment conducted in Portugal

**Nitrogen fertilizer reduced by 20%**

Compared to general farmers' average

## Visualizing global environment and people's behavior

### Capturing real-world data and communicating it by fiber optics

- Collect data in 3D space for use in monitoring places and systems
- Capture data around existing long-distance fiber optics with high sensitivity and map it on 3D space
- Traffic congestion in expressways, anomaly detection in places where people gather, preventive detection of equipment failures at factories, intrusion detection at large-scale facilities, etc.



## Technical challenge

Detecting various state change over a wide area **at low costs and using as little space as possible**

**By applying fiber-optic cables used for communication as sensor media, environmental changes (vibration, temperature, distortion) and its location on the cable can be sensed.**

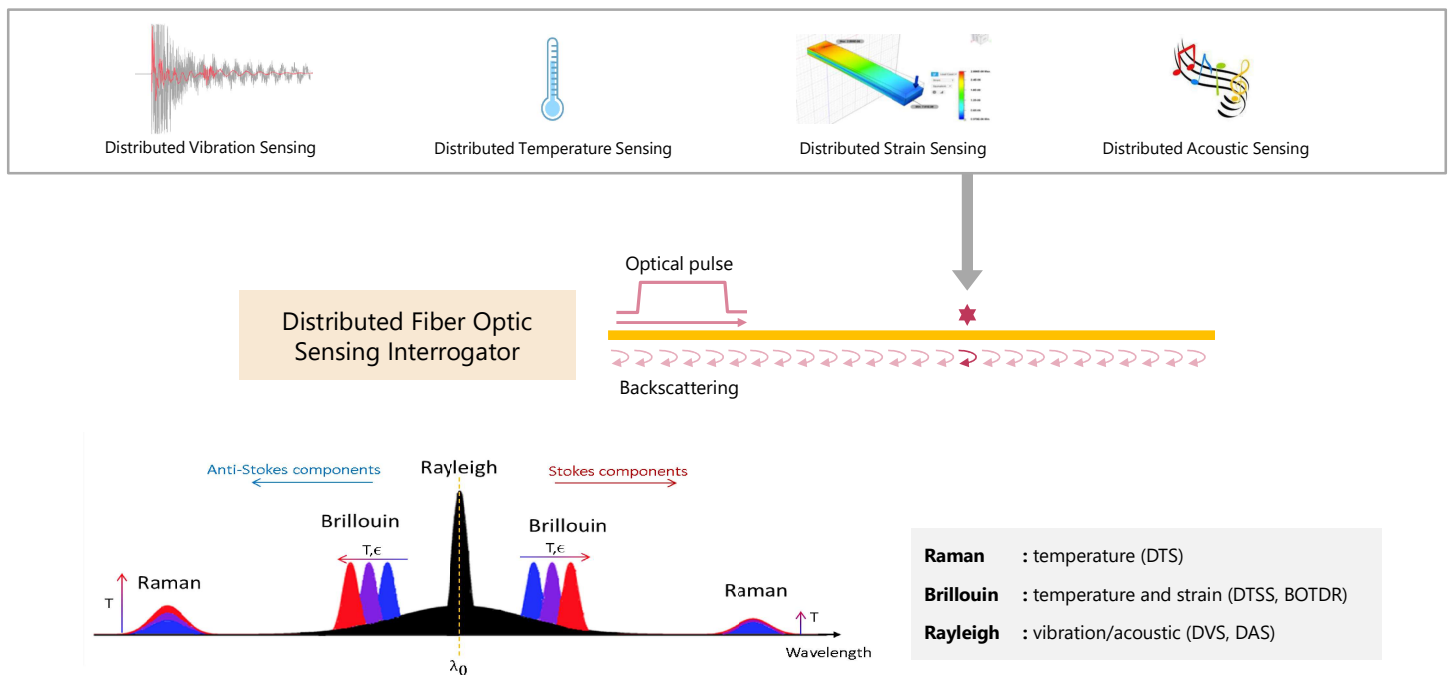


High-coherence sensing signals are input from one end of the fiber-optic cable to observe backscattered light. The changes in backscattered light are used to capture environmental changes, and the location is identified based on the time elapsed from such data input to the point of observation.

Movie



## Visualizing global environment and people's behavior



49

© NEC Corporation 2021

Orchestrating a brighter world **NEC**

## Summary

Advanced ICT for Digital transformation of our social systems

## ◆ For better life

- biometric National ID
- AI assisted drug discovery
- shoes embedded sensor



## ◆ For innovative society

- automated operation of complex infrastructure
- secure data analysis system
- prescriptive analysis and actuation



## ◆ For better environment

- Energy-saving DC cooling
- AI aided farm management
- city-wide environment monitoring



50

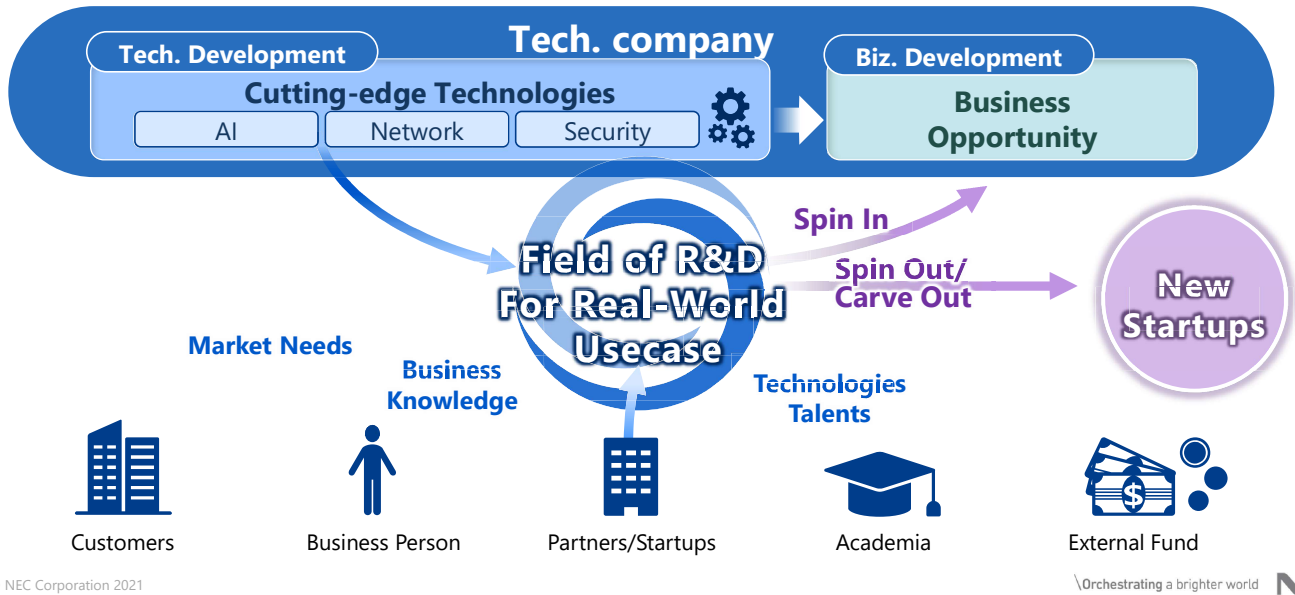
© NEC Corporation 2021

Orchestrating a brighter world **NEC**



## Era of ecosystem-oriented R&D

Expand NEC's technologies externally at an early phase, taking in technologies and funds from customers/startups/VC to speed up R&D. Open innovation of an Inbound/Outbound fusion type



*Let's bring amazing ideas to life, together.*



## Special Thanks to...



### Life

#### Finger print

Y. Koda (Biometrics Res. Labs., NEC)

#### AI Drug discovery

S. Niccolini, M. Niepert (NEC Labs. Europe)  
Y. Yamashita (AI Drug Dev., NEC)

#### Gait analysis

K. Nakahara (Biometrics Res. Labs., NEC)  
R. Omori, A. Furukawa, N. Kawata  
(Corp. Biz. Dev., NEC)



### Society

#### Virtualized RAN

X. Costa (NEC Labs. Europe)

#### Data Sharing PF

S.Fujii, S.Yagi, Y.Etou (Tech. Value Cre., NEC)

#### Robotics

H. Yoshida (System PF Res. Labs., NEC)



### Environment

#### Data Center Cooling

M. Yoshikawa (Biz. Incubation, NEC)

#### Smart Agriculture

N. Ooki (Corp. Biz. Dev., NEC)

#### Optical Fiber Sensing

T. Wang, J. Fang (NEC Labs. America)  
T. Hino (Data Science Res. Labs., NEC)

### Especially for

S. Senda (System PF Res. Labs., NEC)   M. Uekubo (Global Innovation Strategy, NEC)   M. Inafune (Corp. Design., NEC)

## Orchestrating a brighter world

NEC creates the social values of safety, security,  
fairness and efficiency to promote a more sustainable world  
where everyone has the chance to reach their full potential.





Session 7:  
Application I  
( Chair: Tomoo Inoue )



# A System To Directly Feed Back the Audience's Attention Ratio To the Presentation Venue

Yuichi Takyo<sup>†</sup>, Katsuhiko Kaji<sup>‡</sup>

<sup>†</sup>Graduate School of Management Information Science, Aichi Institute of Technology, Japan

<sup>‡</sup>Department of Information Science, Aichi Institute of Technology, Japan

**Abstract** - While presentation skills are becoming more and more important in society due to the need for logical explanatory skills, many working people and students are not good at presentations. In addition, even for those who are used to giving presentations, it becomes difficult to grasp the situation of the audience as the size of the venue and the number of audience members increase. In this study, we propose a presentation system that estimates the entire audience and improves the awareness of both the presenter and the audience. The system estimates the audience's gaze, and uses the attention ratio as an indicator of whether a presentation is good or bad. In order to feed back the attention ratio in real time to the participants of the presentation including the presenter, we present the effect on the slides. We believe that the presenter can infer the state of the presentation from the state of the slides because the effects are presented according to the attention ratio. In order for the presenter to look back on the presentation, the transition of the attention ratio is visualized after the presentation, and the slides are displayed according to the time axis. In the evaluation experiment, we used the two visual effects of this system in an actual presentation and evaluated whether they lead to an improvement in awareness. The results showed that the visual effect of displaying the attention ratio with an increasing/decreasing gauge did not improve the awareness of both the presenter and the audience. On the other hand, the visual effect of changing the lightness and darkness led to a difference in opinion among the presenters and an improvement among the audience.

**Keywords:** Presentation, Groupwear, Interaction, Visual Effect, Line of Sight

## 1 INTRODUCTION

Presentation skills are becoming more and more important as the ability to explain things logically becomes necessary in society. Presentations are used for conferences, product promotion, university classes, and academic presentations. In order to develop presentation skills, students need to be familiar with presentations from an early age. In order to improve their presentation skills, many schools have introduced presentation practice classes for elementary school students.

Many people are uncomfortable with presentations. We believe that some of the problems that contribute to this dislike are caused by the behavior of the presenter and the audience during the presentation. One bad example is that the presenter concentrates on the presentation and does not pay attention to the audience. The presenter is unable to check whether the

audience is paying attention to his or her presentation, and the audience becomes distant or dozes off.

Even for those who are accustomed to giving presentations, such as teachers, there is a need for continuous efforts to improve class-style presentations. Monitors installed at individual desks or at the rear of the classroom allow the audience to listen to the presentation even if they are seated far from the presenter. However, there is a limit to the number of people that the teacher in front can be aware of, and as the audience increases, the teacher becomes less aware of what is going on in the back rows. If the teacher does not understand the situation of the audience and teaches a one-sided lesson, the audience's comprehension may decrease due to strangers or falling asleep.

Therefore, this study proposes a presentation system that aims to improve the awareness of both the presenter and the audience based on the attention ratio of the audience. The outline of the presentation system is shown in Figure 1. As a method to determine whether the audience is interested in the presentation, we focus on the attitude of the audience. Those who are not interested in the presentation often look down due to dawdling or dozing. Using this habit, we estimate the audience's gaze using posture recognition and calculate the attention ratio to the presentation. We present the effect of the attention ratio using a medium that is common to both the presenter and the audience. When the attention ratio becomes low, a negative visual effect is given, and both parties are given a sense of crisis.

## 2 RELATED RESEARCH ON PRESENTATION SUPPORT

One of the most important aspects of giving a presentation is how to make it easy for others to understand. Gestures are a way of conveying information to the audience in an easy-to-understand manner. However, it is difficult for presenters who are not familiar with presentations to convey information using gestures, and they often use only words to convey information. There is a research on detecting a presenter's speaking style and body movements and feeding them back to the presenter using a wearable device worn by the presenter [1][2]. The system detects the presenter's state based on the presenter's speaking speed, intonation, body orientation and movement. Based on the detected information, feedback is provided through Google Glass worn by the presenter. In these studies, feedback is given only to the presenter, and the audience only listens to the presentation. Since no feedback is given to the audience, it does not affect the improvement of



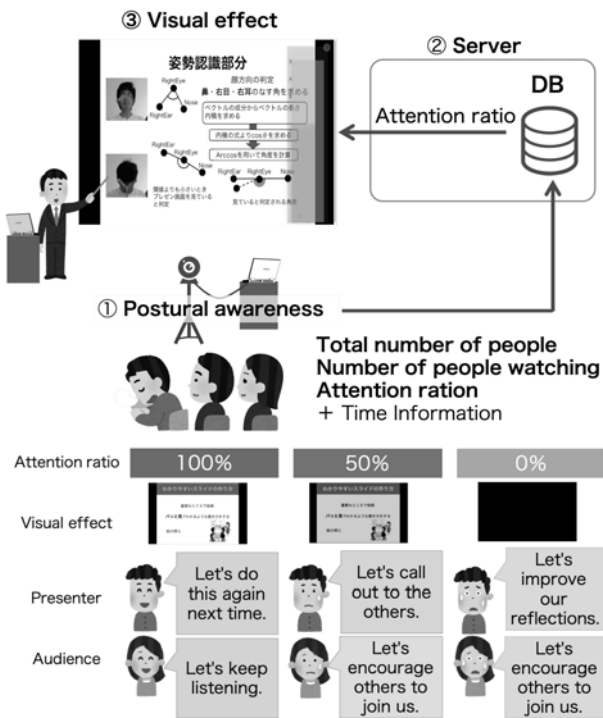


Figure 1: Schematic diagram of the study

the audience's listening attitude.

One of the behaviors of presenters who are inexperienced in presentations is that they only look at their own slides and cannot check the situation of the audience. Without being able to check the situation of the audience, they cannot judge whether their presentation has been understood by the audience, and thus cannot improve their presentation. There is a system in which comments from the audience are shown on the screen during the presentation[3][4]. The audience can express their own impressions and voices to the presenter. In addition, the presenter can respond to some of the comments. There is a system that feeds back real-time audience response to the presenter using AR goggles[5]. The feedback, including emotions and questions, is displayed around the audience who responded. Similarly, there are systems that provide feedback to the presenter using a remote control or a smartphone[6][7][8]. In these systems, devices are used. The devices provide feedback to the presenter. These feedbacks can be used as an indicator to reflect on the presenter's presentation in real time or after the presentation. In these systems, the audience needs to take actions to give feedback to the presenter, so they cannot concentrate on listening to the presentation.

There is a research on estimating the degree of interest and concern of a seated person from a pressure sensor attached to a chair[9][10]. The posture of the person sitting in the chair is estimated from the values obtained from the pressure sensor. Based on the estimated posture, the degree of interest is measured. In these studies, multiple pressure sensors are installed on the seat and back of the chair. The cost of purchasing and installing the sensors is high.

### 3 PRESENTATION SYSTEM "slipe"

We propose a presentation system called "slipe", which can influence both the presenter and the audience in real time by presenting effects according to the attention ratio to a common medium. The term "slipe" was coined from the initial letters of system, listener, incite, presenter, and elevate.

#### 3.1 System Configuration

The configuration of the system is divided into posture recognition and video effects. The situation of the audience is input to a PC for posture recognition using a Web camera. The system estimates the audience's gaze and field of view from the posture recognition, and judges whether they are watching the presentation or not for each individual. The judgment is made for the number of people detected by the Web camera. From the results, the attention ratio of the audience to the presentation is calculated. The total number of people detected, the number of people watching, and the calculated attention ratio are sent to the server. The server stores the total number of people, the number of people watching, the attention ratio, and the time of the transmission.

In this system, the slides projected by the projector are used as a medium to influence both the presenter and the audience. Fig. 2 shows the presentation situation in which the visual effects are presented. When the attention ratio is low, a visual effect that gives a sense of crisis to both parties is presented. When the attention ratio is high, we present a visual effect that encourages both parties to maintain this state. We aim to improve the awareness of both the presenter and the audience through the presented effects. For example, when the attention ratio of the presentation is low, the visual effect shown by the red dotted line in Fig. 2 is presented. If the presented visual effect indicates that the attention ratio is low, the presenter should try to call the audience's attention. The attention ratio saved on the server can be checked after the presentation and used to reflect on one's own presentation.



Figure 2: Presentation of the visual effect presented

#### 3.2 Calculate the Attention Ratio

In this system, we estimate the audience's gaze and use the attention ratio as an indicator of whether a presentation is

good or bad. The attention ratio is the ratio of the number of people who pay attention to the presentation to the number of people in the audience. Special devices such as eye tracking and leg estimation can be used to accurately estimate whether the audience is paying attention to the presentation. However, these devices are costly because they need to be prepared for each person in the audience. Another method is to provide feedback from the audience using a smartphone or a remote control. This method requires the audience to perform some operation, which makes it difficult for them to concentrate only on the presentation. In our system, a single camera is used to estimate the entire audience. Since the status of the audience is estimated using a camera, the audience does not need to perform any operation and can concentrate on the presentation.

In this system, we use OpenPose to estimate the audience's gaze and calculate the attention ratio by using the feature points obtained from posture recognition. OpenPose is a framework for estimating multiple human feature points in real time from a single image presented by Carnegie Mellon University (CMU) in the United States, and it estimates each feature point for each person. There are 25 feature points to be estimated, and our system uses 6 of them: nose, ears, eyes and neck. The gaze of each individual is estimated as a vector from the feature points of the face and neck obtained by OpenPose. The field of view and the position of the slide are set as rectangles for each individual in the audience. The gaze is defined as a vector from the midpoint of the detected feature points of both ears to the feature point of the nose. In the case where the subject is looking sideways, only one of the ear feature points can be detected, so the vector from that feature point to the nose is used as the line of sight. The field of view of each individual set by the gaze vector is expanded by perspective projection. As shown in Fig.3, if the expanded field of view is partly within the range of the slide, the individual is paying attention to the presentation. This process is performed for the number of people detected, and the number of people detected, the number of people watching, and the calculated attention ratio are sent to the server.

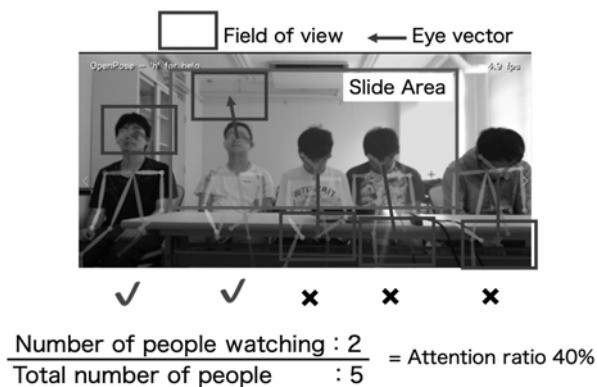


Figure 3: Determination of attention by estimated field of view and calculation of attention ratio

### 3.3 Presentation of Visual Effects

In order to feed back the attention ratio in real time to the participants of the presentation including the presenter, we present the visual effect to the slides. Since the visual effects are presented on the slides according to the calculated attention ratio, we believe that the presenter can infer the state of the presentation from the state of the slides. In the same way, the audience can infer the state of the presentation from the visual effects presented on the slides they are looking at, so they can call out to those around them and contribute to improving the quality of the presentation. There are two types of visual effects presented on a slide: one is a gauge that increases or decreases the attention ratio, and the other is a change in brightness or darkness.

Fig. 4 shows the visual effects presented by changes in the attention ratio. Normally, only the presenter needs to know the status of the audience such as the attention ratio, but in our system, the attention ratio is also presented to the audience. The visual effect of displaying the attention ratio as a gauge (Fig. 4) is that if the presented attention ratio is low, the displayed gauge turns red, which makes the presenter feel threatened. On the other hand, when the attention ratio is high, the gauge turns green, which indicates that the presenter's presentation is going well. In addition, since the audience can see the attention ratio of the entire audience, we expect this to have the effect of encouraging them to improve their awareness so that they can concentrate more on the presentation.

On the other hand, the visual effect of changing light and dark (Figure 5) is a negative visual effect that makes it difficult to continue the presentation. The lower the attention ratio, the darker the image becomes, and the more critical it is for both the presenter and the audience. It will return to normal brightness as the attention ratio increases. Especially for the presenter, if the quality of the presentation is poor, the screen will become dark and it will be difficult to continue the presentation. Therefore, in order to continue the presentation, the presenter is expected to improve the presentation expression method in real time by devising an expression on the spot and calling out to others. However, when the attention ratio is high, the presentation is the same as a normal presentation, so it cannot have a positive impact.

### 3.4 Visualization of Changes in Attention Ratio

In order for the presenter to reflect on the presentation, the system visualizes the transition of the attention ratio after the presentation and displays the slides according to the time axis. This system provides real-time feedback by showing the visual effects on the slides during the presentation. However, it is not possible to reflect on the entire presentation only by presenting the visual effects. For this reason, the attention ratio calculated during the presentation is stored on the server along with the calculated time. The start and end of the presentation are also managed in the server, and the transition of the attention ratio is presented in the presentation at the end of the presentation. In addition, Microsoft PowerPoint has a

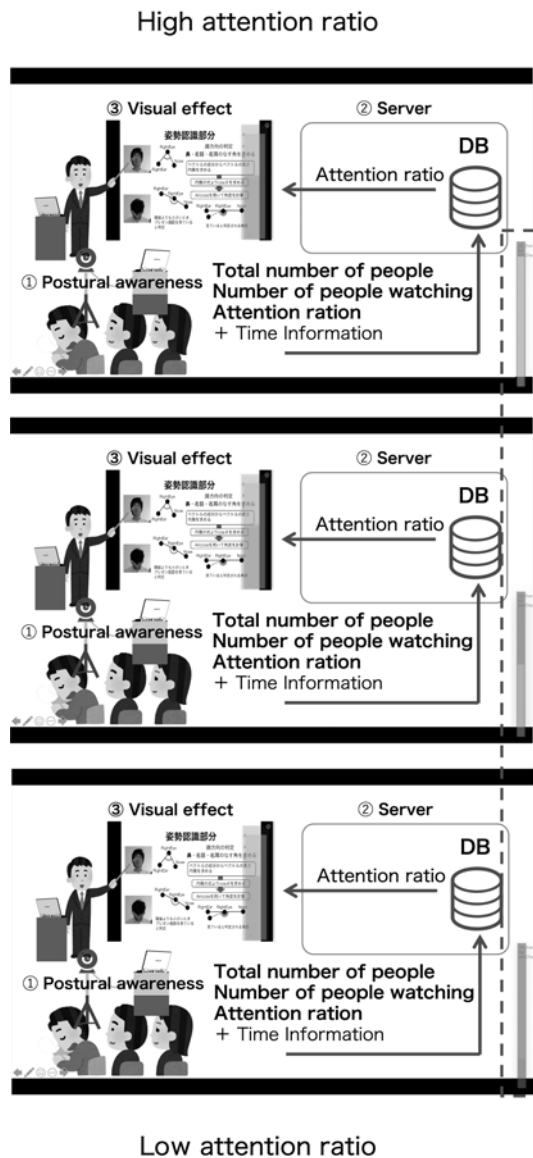


Figure 4: The visual effect of presenting the attention ratio as a gauge

function to record the display time of slides. The recorded time and the slides used in the presentation can be output. The output slides and time are displayed together with the transition of the attention ratio of the presentation. Since the attention ratio for each slide is visually clear, it is easy to understand which slide is the cause of the increase or decrease in the attention ratio. In addition, comparing the changes in the attention ratio with the displayed slides will be useful for improving the presentation method and creating future slides.

## 4 EVALUATION EXPERIMEN

We conducted an evaluation experiment to see if this system could improve the awareness of the presenter and the audience. Four presenters gave presentations using this system to an audience of six people. The system presented the gauge and light/dark change as visual effects. A total of eight pre-

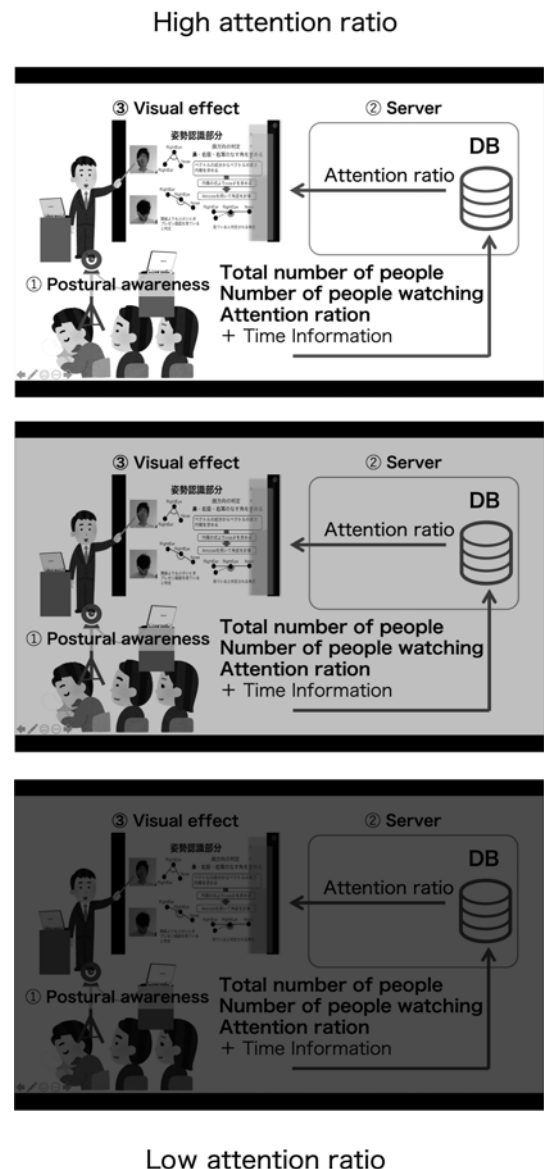


Figure 5: Effects of lightness and darkness depending on attention ratio

sentations were made, four each of the visual effects of displaying the attention ratio and the visual effects of light/dark changes. Afterwards, a questionnaire was given to both the presenters and the audience. The presenters were asked to respond to the following questions on a 5-point scale: "Whether the feedback during the presentation led to improved speaking awareness," "Whether this system clarified areas for improvement in my presentation," and "Whether the feedback after the presentation led to future improvements". The audience was also asked to rate on a 5-point scale: "Whether the feedback during the presentation improved their listening attitudes", and "Whether the feedback during the presentation interfered with their listening". The average of the questionnaire results of the presenters and the audience is shown in Figure 7. Q1 to Q5 in Figure 7 indicate the following questions.

[Q1 ] Whether the feedback during the presentation led to

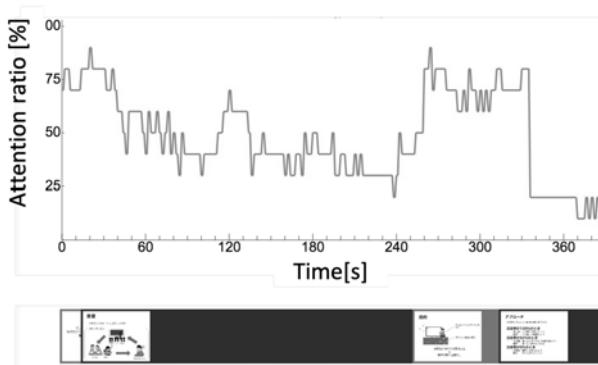


Figure 6: Visualization of attention ratio and slides

improved speaking awareness

[Q2 ] Whether this system clarified areas for improvement in my presentation

[Q3 ] Whether the feedback after the presentation led to future improvements

[Q4 ] Whether the feedback during the presentation improved their listening attitudes

[Q5 ] Whether the feedback during the presentation interfered with their listening

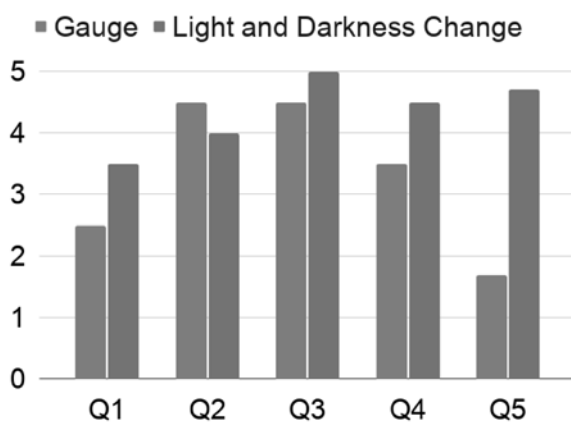


Figure 7: Average of survey results

As for the presenters' feedback during the presentation, the visual effect of changing the lightness and darkness according to the attention ratio was higher than the visual effect of displaying with a gauge in terms of improving their awareness of speaking. In addition, both the light/dark visual effects and the gauge visual effect clarified the improvement points of the presenter's presentation, and the feedback after the presentation led to future improvement. In the case of the audience, whether the feedback during the presentation led to an improvement in their listening awareness or hindered them from listening to the presentation, both the light/dark visual effects were higher. In particular, the light/dark visual effect was found to be a hindrance to the presentation.

## 5 CONSIDERATION

The visual effect of displaying the attention ratio as an increasing/decreasing gauge did not improve the awareness of the presenter during the presentation. It is thought that the presenters were concentrating on their own presentations and did not have time to look at the displayed gauges. In contrast, a slight improvement in the awareness of the visual effect of changing light/dark was expected. This may be because the presentation slides darkened during the presentation, which made it difficult to proceed. One presenter forced himself to proceed with the presentation using the presenter's tool displayed on the PC at hand, even though the displayed slides were darkened by the visual effect. Another presenter tried to call out to the people around him when the slides were darkened, but he answered that he did not know what to say. It can be said that only presenting the visual effect from the system does not have a useful influence on the presenter.

For the audience, we could not improve their awareness during the presentation by displaying a gauge that increased or decreased the attention ratio. This is because the audience concentrated on the presentation and did not pay attention to the gauge, but only watched the increase and decrease of the attention ratio. On the other hand, the visual effect of varying light and dark led to an increase in the audience's awareness of listening. When the slides went dark, some participants wanted to look around to see how many people in the audience were not watching the presentation. When they tried to talk to the people around them, they found it possible to talk to acquaintances, but difficult to talk to strangers. We believe that the temporary pause in the presentation due to the darkening of the slides led to an improvement in listening awareness. In this study, the visual effect of the darkening of the slides was shown to the audience not only on one part of the slides but also on the whole slides.

The feedback after the presentation indicated that it would lead to future improvements. For the visual effect of displaying the increase or decrease of the attention ratio as a gauge, the presenter did not have time to look at the gauge. On the other hand, the screen darkened during the presentation in the visual effect of changing lightness and darkness. Therefore, it can be said that the presenters were able to recognize what points needed to be improved during the presentation by comparing the presentation slides and the changes in the attention ratio. In addition, it was answered that the changes in the presentation slides and the attention ratio clarified the scenes in which the audience was interested and those in which they were not interested.

## 6 CONCLUSION

In this paper, we propose a presentation system that estimates the entire audience and improves the awareness of both the presenter and the audience. There have been proposals to present some kind of feedback to the presenter, but they require manipulation of devices, and the audience cannot concentrate on the presentation. To solve this problem, we proposed a presentation system that estimates the gaze of each individual audience member based on posture recogni-

tion, presents visual effects based on the attention ratio, and provides feedback in real time without the need for the audience to operate a device. In order to feed back the attention ratio in real time to the participants of the presentation including the presenter, the visual effects are presented on the slides. In addition, we compared the change of the attention ratio and the slides used in the presentation as feedback after the presentation. In the evaluation experiment, we used the two visual effects of this system in an actual presentation and evaluated whether they lead to an improvement in awareness. The results showed that the visual effect of displaying the attention ratio with an increasing/decreasing gauge did not improve the awareness of both the presenter and the audience. On the other hand, the visual effect of changing the lightness and darkness led to a difference in opinion among the presenters and an improvement among the audience.

In the future, since the number of evaluation experiments is very small (4), we will continue to conduct evaluation experiments by using this system at any time during presentations. In addition, we will implement and evaluate visual effects that affect both the presenter and the audience.

## REFERENCES

- [1] I. Damian, C. S. (Sean) Tan, Tobias B., Johannes S., Kris L., and E. André. Augmenting social interactions: Real-time behavioural feedback using social signal processing techniques. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, CHI '15*, p. 565–574, New York, NY, USA, 2015. Association for Computing Machinery.
- [2] M. I. Tanveer, Emy L., and Mohammed (Ehsan) H. Rhema: A real-time in-situ intelligent interface to help people with public speaking. In *Proceedings of the 20th International Conference on Intelligent User Interfaces, IUI '15*, p. 286–295, New York, NY, USA, 2015. Association for Computing Machinery.
- [3] Commnet screen. <https://commentsscreen.com/>, 6 2021.
- [4] Niconico gakkai β. <http://niconicogakkai.jp/info/about>, 6 2021.
- [5] Dhaval P. and Timothy B. Making it personal: Addressing individual audience members in oral presentations using augmented reality. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, Vol. 4, No. 2, June 2020.
- [6] Robin H. Kay and Ann L. A strategic assessment of audience response systems used in higher education. *Australasian Journal of Educational Technology*, Vol. 25, No. 2, May 2009.
- [7] Vasileios T. and Cesare P. Asq: interactive web presentations for hybrid moocs. In *WWW (Companion Volume)*, pp. 209–210, 2013.
- [8] Jae H. H. and Adam F. Understanding the effects of professors' pedagogical development with clicker assessment and feedback technologies and the impact on students' engagement and learning in higher education. *Comput. Educ.*, Vol. 65, p. 64–76, July 2013.
- [9] Selene M. and Rosalind W. P. Automated posture analysis for detecting learner's interest level. In *2003 Conference on Computer Vision and Pattern Recognition Workshop*, Vol. 5, pp. 49–49, 2003.
- [10] Ricardo B., Ángel P. de M., and Jesús G. B. Development of an inexpensive sensor network for recognition of sitting posture. *Int. J. Distrib. Sen. Netw.*, Vol. 2015, , January 2015.

# Software Edutainment Systems and Analysis of Learners' Data based Docker and Edutainment

Ryosuke Tsutsumi<sup>†</sup>, Wei JiuJun<sup>†</sup>, Shinpei Ogata<sup>‡</sup>, Masaaki Niimura<sup>‡</sup>, Kozo Okano<sup>‡</sup>

<sup>†</sup>Shinshu Graduate School of Science and Technology, Japan

<sup>‡</sup>Faculty of Engineering, Shinshu University, Japan

{20w2061k, 20w2030k}@shinshu-u.ac.jp

{ogata, niimura, okano}@cs.shinshu-u.ac.jp

**Abstract** - In recent years, it has become necessary to improve the quality of programming education in universities in order to educate 'future' engineers who realize a smart society. The idea of embedding educational elements in entertainment such as games, called Edutainment, can be also used in programming education to provide highly effective learning opportunities. This approach can be realized by constructing a programming practice environment with game elements called Gamification, and introducing it to programming education in university classes. In this study, we constructed a remote educational environment using code-server as an IDE, and applied it to actual programming exercises. In addition to assigning thirteen types of programming tasks to students, we collected students' data such as source code and working history data for 13 students, and obtained log files of the number of files for each task data, the final capacity, and the change in capacity compared to the initial value. In this paper, we report the outline of the system and the application of the exercise, as well as the data analysis and discussion based on the obtained data.

**Keywords:** Software engineering, Java, Edutainment, Gamification, Docker

## 1 INTRODUCTION

Education using digital games has been of interest since the beginning of the spread of home computers, and various efforts have been made over the years. In addition, as information and communication technologies have evolved, various studies have been conducted, and many findings and issues have been presented. In particular, in recent years, research on educational methods using digital games has become active[1][2][3], especially overseas.

Currently, although there are Edutainment approaches to programming education in Japan[28], there are few of them in higher education institutions such as universities. Previous studies[4] have reported on the implementation of programming exercises by building a virtual environment on Linux using an open source software called Docker[5][15]. In addition, there have been several approaches to providing educational environments using Docker in the past[6][11][12]. From these approaches, we determined that Docker is effective in providing an educational environment for a large number of people. In addition, by using a cloud IDE called code-server[22], we believe that the prepared environment can be used in remote classes[21]. In other words, the final goal of this study is to verify the learning effectiveness of Java programming exer-

cises by using Edutainment and Gamification approaches in programming education for university students.

In this report, we summarize the results of our investigation of the definition of the difficulty level of the tasks and the students' commitment to the tasks, based on the task data and operation history data obtained through the Java programming exercises.

In Section 2, we describe the knowledge required for this report, and in Section 3, we describe the configuration of the educational support system we used. Section 4 gives an overview of the data collected in the experiment. Section 5 describes the analysis using the obtained data, Section 6 discusses the results of the analysis. Final, Section 7 gives a summary.

## 2 PRELIMINARIES

### 2.1 Edutainment and Gamification

Edutainment is notion for methods of increasing interest in learning through the appeal of games[13]. This concept was first proposed in the late 1980s and 1990s. The development of multimedia educational software during this period led to the spread of the concept of "Edutainment" and the introduction of learning games into school education and home education. Furthermore, in the 2000s, digital game technology becomes more advanced, making it possible to use them at low cost. This has increased interest in the potential of games in other fields[10]. In the 2010s, the concept of "Gamification"[17] began to be proposed, which aims to improve the motivation and performance of users when they engage in a task by incorporating game mechanics into the task.

The historical evolution of education using digital games is shown in Figure 1.

### 2.2 Docker

Docker is a platform for building an application execution environment by creating a container based on a single image that contains the code, runtime, and system tools, required to run an application[16]. Docker is characterized by extremely fast startup and shutdown[14] because only the userland is virtualized. In waterfall application development, system development proceeds in the following order: development, test, staging, and production environment. Therefore, even if the system works correctly in the development and test environments, it may not work properly when deployed in the subsequent environments. A staging environment is a

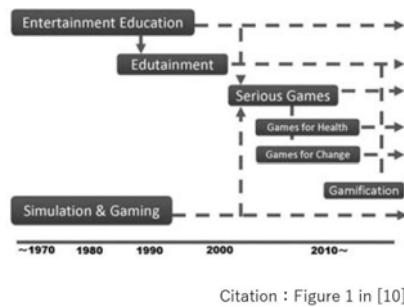


Figure 1: The historical evolution of education using digital games

test environment that is checked just before deploying a developed application to a production environment in system development where continuous delivery is performed.

Docker manages these infrastructure environments as containers. In other words, all the files and directories required to run an application are organized as a container. The Docker images that form the basis of these containers can be shared in repositories such as Docker Hub. The developer then uses Docker to create a Docker image that contains everything necessary to run the developed application. This image serves as a template for the container, and the container is run based on it. This image can be run in basically any environment where Docker is installed, reducing the risk that the image will work in the development and testing environment but not in the production environment.

### 2.3 The IDE code-server

An IDE code-server is a cloud IDE version of Visual Studio Code [22]. Using code-server avoids the port number problem mentioned in Section 3. Therefore, it can be accessed from a browser to work remotely, and Visual Studio Code plug-ins can also be installed.

### 2.4 Code Testing

A process called “code testing” is required to ensure that a program works properly [18]. Of the code testing, the testing that takes into account detailed conditional branching within the program is called white box testing [23]. Users can improve the quality of the programs they created by performing tasks with accomplishing white box testing in mind. In addition, the ratio of the portion of the program that has been tested is called code coverage. The following are the main steps to measure this.

First, in statement coverage[25], all instructions in the program are tested to ensure that they are executed. Then, in branch coverage[26], is tested to ensure that all branching operations in the program are performed. And finally, in condition coverage[24], tests to ensure that all the results of the conditional expressions in the program are satisfied. By the

way, branch coverage is not considered in condition coverage, so even if the latter is satisfied, the former may not be satisfied.

## 3 THE STRUCTURE OF EXERCISE SYSTEM

In the previous study[4], a containerized virtual environment using Linux and Docker was constructed as shown in the Figure 2. In addition, the course management system for online learning called Moodle[29], and an IDE called Eclipse Che were used to provide users with a programming practice environment and enable the collection and storage of users’ learning history.

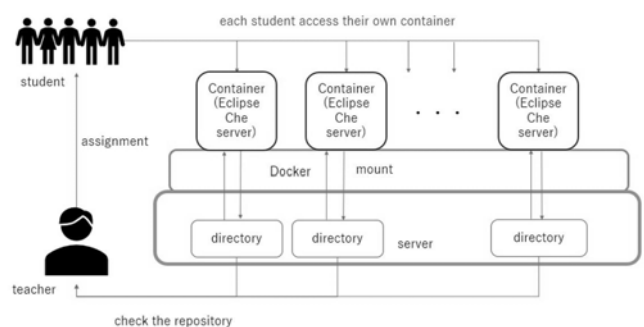


Figure 2: The environment used in the previous study

Creating containers by using this environment, and sharing images on a repository service called Docker Hub. This image allows the students to use the images created from the containers.

In addition, we created a container in the Docker image like in Figure 3, which can be linked to each student’s student ID and the student. This system links the user’s student ID with the port number for accessing the environment. This system allows each of users to automatically access each exercise container when they access the university’s network service.

Furthermore, we used code-server[22] instead of Eclipse Che as the IDE. The reason for this is due to the port number used in Docker. Docker uses 8080 port for code-server and 32000 or higher port for Eclipse Che. However, when using the university network, it is not possible to use a port number with a large value due to safety issues. For the same reason, functions such as in Figure 3 cannot be used remotely in Eclipse Che. Therefore, we thought that using code-server would be suitable for building a remote environment.

As a result, what we have provided in this study is an exercise environment as shown in Figure 4, which allows each user to work in a remote environment.

## 4 DATA COLLECTING

In the experiment conducted in the previous study, we succeeded in recording and collecting students’ data such as



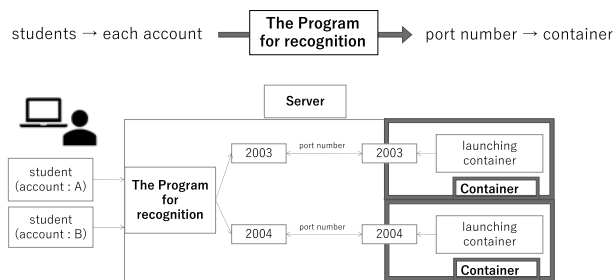


Figure 3: Linking users and each port number

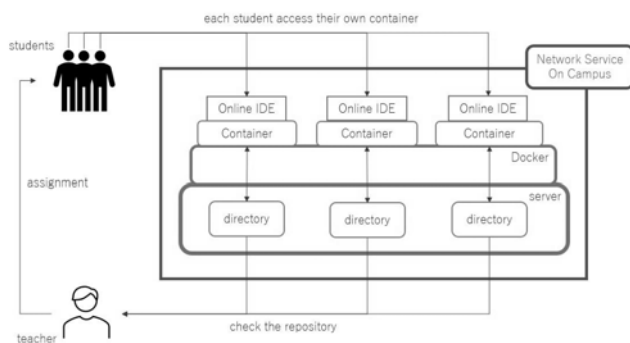


Figure 4: The environment used in this study

source code and working history data. The following information are the details of the targeted exercises.

**The period of the exercise** from October 30, 2020 to January 22, 2021

**The class name** Software Engineering

**The target users** Seventy-three third-year undergraduate students in the Faculty of Engineering at Shinshu University who attended the above class

**The Contents of the exercise** Java programming exercises (ten basic tasks, three advanced tasks)

**The Periods to be collected** Every 1.5 seconds, and when there is a change from the previous data

The following elements were collected in the exercise.

1. Programs created in the exercise
2. Test cases created in the exercise
3. Users' operation history
4. Access date and time of each user for each exercise

## 5. Cumulative access time of each user for each exercise

In the previous year(2019), two 90-minute exercises were conducted in a row, but this year's lecture was conducted in an online format.

### 4.1 Learners' Data of Exercises

In the exercise, the students were asked to create a total of 13 test cases, ten basic and three application test cases in Java.

There are several tools for creating test cases[19][20], but in this study, we used JUnit can be used in code-cover. JUnit is a framework[18] for automatically executing Java tests.

By verifying the above files, we research the progress of the task by each user.

### 4.2 Test case

In the test case creation task that users do in this study, two types of tasks are required to be satisfied: branch coverage and condition coverage.

### 4.3 Users' operation history

We have also implemented a function to collect the user's operation history to ensure the security and recoverability of the users' task data. In addition to checking the issue file every 1.5 seconds and collecting only the changes, we also collect the submitted issue file, i.e. the last version.

### 4.4 Access log

It also collects the date and time of user access and the total access time. In this report, we only use the case of Basic Task 6 as an example.

### 4.5 Elements included in the data

The data to be collected includes the following elements

- The time and date of the exercise
- The change in the size of the task file
- The number of accesses
- The total number of files collected
- Access date for each student
- The total access time for each student

By verifying these results for each task and each student, it is possible to consider the students' approach to the task and their progress.

Table 1: The number of files for each task data

	Student 1	Student 2	Student 3	Student 4	Student 5	Student 6	Student 7
btask1	11	22	7	11	9	28	5
btask2	108	113	96	36	98	63	61
btask3	113	83	131	511	62	180	154
btask4	323	118	268	642	2	191	282
atask1	623	201	464	1,003	430	283	241
btask5	65	79	59	161	45	9	104
btask6	465	222	577	263	540	633	290
btask7	87	40	89	287	128	81	75
btask8							
atask2	671	423	492	507	366	694	120
btask9	56	52	292	63	142	104	119
btask10							
atask3	1,186	996	1,510	2,219		1,107	1,155

	Student 8	Student 9	Student 10	Student 11	Student 12	Student 13
btask1	44	5	6	12	31	64
btask2	60	24	33	173	95	154
btask3	82	49	69	91	493	365
btask4	230	59	100	129	263	361
atask1	12	367	156	205	668	438
btask5	175	183	63	61	2	124
btask6	50	201	324	302	7	738
btask7	59	43	50	50	134	1,076
btask8						
atask2	700	291	252	319	823	763
btask9	130	50	53	74	32	65
btask10						
atask3	1,992	1,071	634	1,499	1,778	2,761

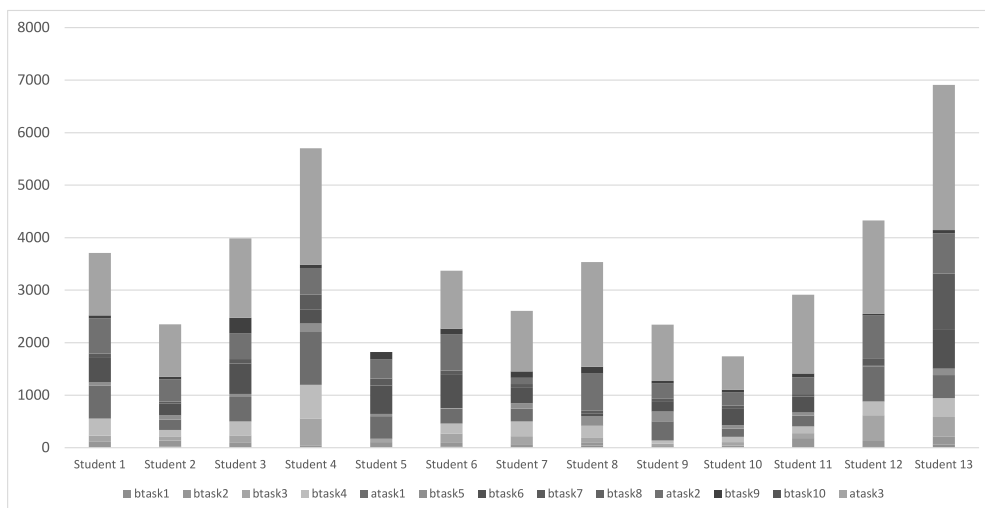


Figure 5: The number of files for each task data

Table 2: Size of each task data (KB)

	Student 1	Student 2	Student 3	Student 4	Student 5	Student 6	Student 7
btask1	4.59	9.26	2.98	3.03	3.86	13.0	2.09
btask2	87.0	105	80.4	26.0	72.0	44.7	48.1
btask3	106	92.9	135	783	59.9	179	189
btask4	371	121	296	597	2.79	214	321
atask1	674	205	468	1340	416	258	172
btask5	61.8	68.0	51.2	198	28.0	10.5	163
btask6	1,540	825	2,000	990	1,930	2,340	958
btask7	64.2	22.3	40.7	195	76.0	50.9	47.2
btask8							
atask2	780	620	564	899	569	884	190
btask9	64.6	55.0	378	64.9	129	110	114
btask10							
atask3	3,000	3,170	5,690	7,870		2,850	3,410

	Student 8	Student 9	Student 10	Student 11	Student 12	Student 13
btask1	19.5	2.11	2.66	5.25	13.6	28.8
btask2	46.0	18.1	33	136	78.2	123
btask3	85.7	57.5	73.2	92.6	581	562
btask4	263	62.7	104	119	248	349
atask1	3.20	458	143	141	470	483
btask5	245	183	41.9	44.5	1.83	104
btask6	223	704	1,190	1,000	29.2	2,520
btask7	40.7	26.3	31.2	30.3	61	650
btask8						
atask2	332	336	398	1,260	1,000	
btask9	132	61.2	60.6	78.2	36.1	75.3
btask10						
atask3	5,370	2,040	2,120	4,270	4,470	7,830

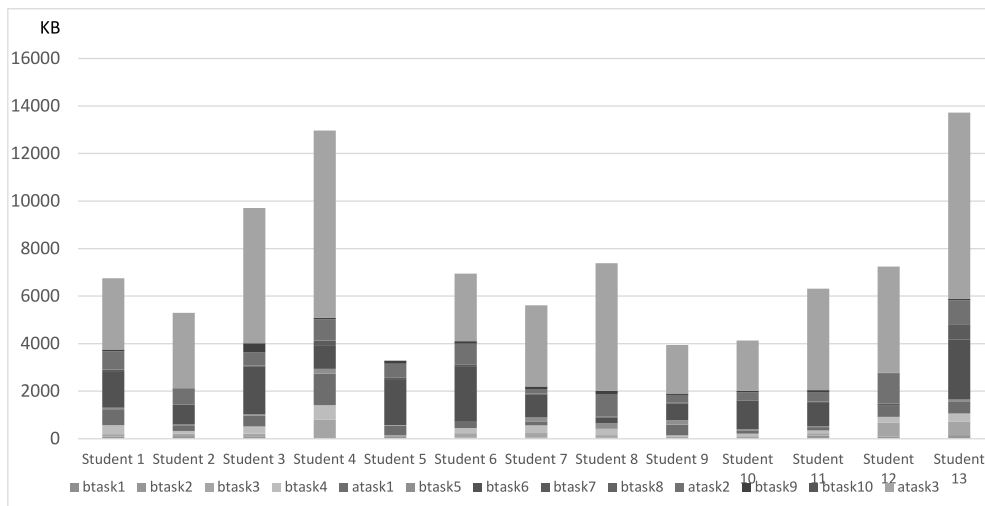


Figure 6: Size of each task data (KB)

Table 3: Size changes of each task data (KB)

	Student 1	Student 2	Student 3	Student 4	Student 5	Student 6	Student 7
btask1	4.59	9.26	2.98	3.03	3.86	13	2.09
btask2	91.5	114	83.4	29	75.8	57.7	50.2
btask3	197	207	219	812	135	237	240
btask4	569	328	516	1,370	138	451	561
atask1	1,210	534	984	2,720	555	709	734
btask5	1,270	602	1030	2920	583	720	896
btask6	2,750	1,320	2,960	3,710	2,470	3,030	1,650
btask7	2,820	1,340	3,000	3,900	2,550	3,080	1,690
btask8							
atask2	3,410	1,810	3,420	4,400	3,110	3,750	1,880
btask9	3,480	1,860	3,790	4,460	3,230	3,850	1,990
btask10							
atask3	6,480	5,040	9,480	1,230		6,710	5,410

	Student 8	Student 9	Student 10	Student 11	Student 12	Student 13
btask1	19.5	2.11	2.66	5.25	13.6	28.8
btask2	65.6	20.2	36	141	92	151
btask3	151	77.7	110	233	677	733
btask4	415	140	243	352	925	1,080
atask1	418	599	399	494	1399	1,560
btask5	663	782	440	540	1401	1,660
btask6	889	1,270	1,640	1,560	1,410	4,180
btask7	930	1,300	1,650	1,590	1,420	4,790
btask8						
atask2	1,430	1,530	1,980	1,990	2,660	5,790
btask9	1,560	1,590	1,990	2,080	2,690	6,030
btask10						
atask3	6,940	3,630	4,100	6,360	7,510	1,386

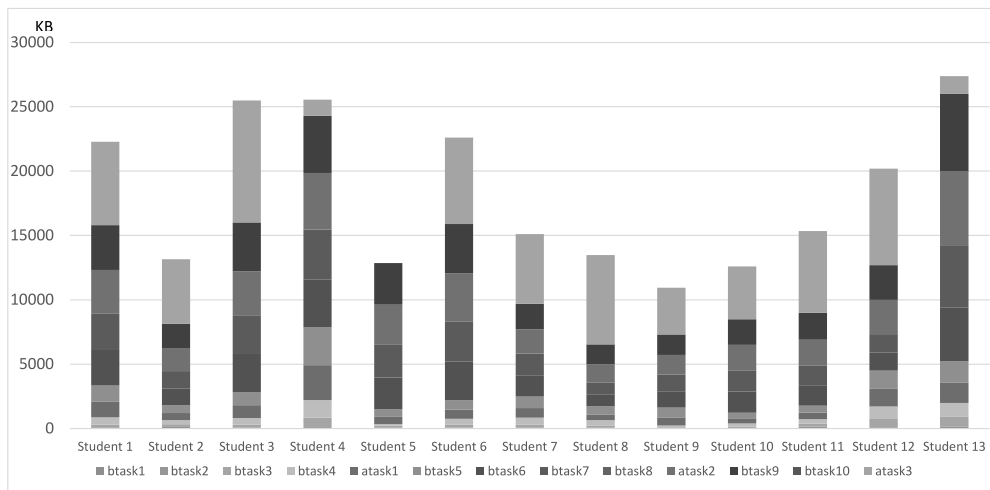


Figure 7: Size changes of each task data (KB)

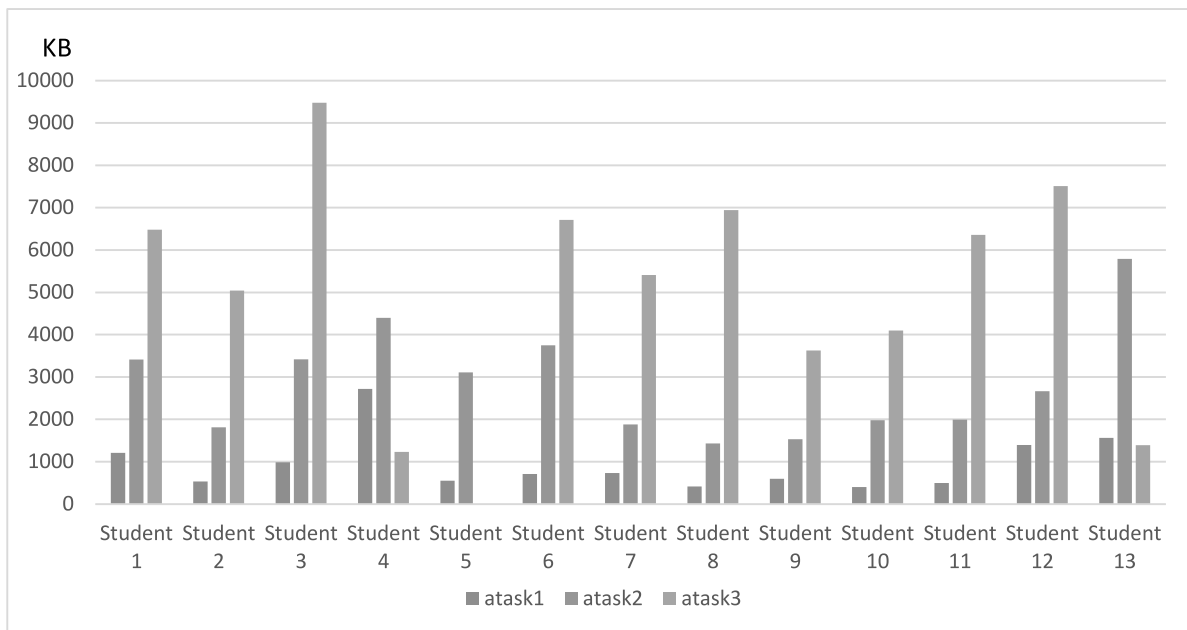


Figure 8: Changes in the size of task data in three advanced tasks

Table 4: The date of student access and the total access time for 'btask6'

		access date (December 2020)						total access time (minute)
Student 1	15th, 7:49:00	15th, 10:55:00	15th, 16:23:00	15th, 19:35:00	16th, 3:58:00	16th, 10:50:00		255
Student 2	14th, 6:27:00	14th, 20:07:00	17th, 12:01:00					130
Student 3	13th, 8:59:00	14th, 11:03:00	14th, 16:17:00	15th, 7:44:00	17th, 12:25:00			420
Student 4	16th, 14:18:00	17th, 5:11:00	17th, 13:28:00					240
Student 5	17th, 6:06:00							210
Student 6	16th, 9:26:00							195
Student 7	16th, 11:28:00	16th, 12:47:00	17th, 3:12:00					190
Student 8	17th, 9:54:00							290
Student 9	16th, 16:48:00							58
Student 10	13th, 18:34:00	13th, 21:44:00	17th, 19:25:00	17th, 20:50:00				200
Student 11	17th, 12:14:00	17th, 13:39:00	17th, 14:37:00					120
Student 12	17th, 05:09:00							4
Student 13	17th, 05:06:00	17th, 14:32:00						265

## 5 DATA

Table 1 and Figure 5 shows the number of assignment data files obtained. Table 2 and Figure 6 shows the final size of each assignment file. Table 3 Figure 7 shows the change in size of each assignment file compared to its initial value. Table 4 shows the access date and cumulative access time for each student in Basic Task 6. Figure 8 shows the size change data

for three questions in the advanced task.

In these figures and tables, “basic task” is referred to as “btask” and “advanced task” as “atask.”

Table 1 shows that the number of files increases for the advanced tasks and the later tasks, i.e., users may have edited a lot of source codes. In addition, the number of files is huge because the files are collected every 1.5 seconds.

Overall, in Table 2, it can be seen that the file sizes are large

for basic task 6 and advanced task 3.

In Table 3, it can be read that the later the assignment is, the more the users are likely to change the code of the assignment to be changed, including the advanced task 1.

Even if we compare only the three advanced tasks, we can notice that the change in size is gradually increasing.

Even if we only check at Table 4, we can notice that there is a large variation in the time students work on the assignment and the access time.

## 6 DISCUSSION

The number of files in Table 1 increases as the user edits the source code. For example, looking at the ‘btask4,’ we can see that Student 5 edited the code only 2 times, while the other twelve students edited the code at least 59 times. This pattern of editing far fewer times than the others can also be applied to Student 6 and Student 12 in ‘btask5’ and Student 12 in ‘btask6.’ This suggests that there are two possibilities: these students were not able to accomplish each task satisfactorily, or they were able to finish their tasks with fewer times.

The size of the task data in Table 2 corresponds to the number of files shown in Table 1. However, when we focus on ‘btask1,’ we see that both Student 1 and Student 4 edited their files 11 times, but there is a small difference in the size of the task data. The following are two possible reasons for this difference, thus it is difficult to judge the correct answer only by the difference in data size.

1. Differences in the length of the source code written by each student
2. The quantity of comment outs

Let’s focus on Student 1 and 8 in Table 4. The difference in the total access time between these two students is about 35 minutes, but there is a large difference in the number of times they access the site. Based on this, if we look at Tables 1, 2, and 3, we can notice that the value for Student 1 is larger overall. From this, we can infer that Student 8 was able to accomplish ‘btask6’ relatively well compared to Student 1. In other words, maybe the first of the above reason is correct about the difference between the number of times of editing files and the amount of data.

## 7 CONCLUSION

We have collected data on tasks created by students in programming classes, their operation histories, and access times. As a result, we found that it is possible to estimate the degree of effort each user puts into each task and the difficulty of the task to some extent. As we collect more data in the future, we will be able to further deepen the relationship between students and tasks.

Future work includes: first, analysis of the obtained data in terms of classifying the tasks and defining the difficulty as a numerical value; second, visualization of the data; and implementation of a gamification function.

## ACKNOWLEDGEMENT

This research was funded by the “Grand Challenge Research for Life Design Innovation (iLDi) Project” run by Osaka University under the “Support for Research Centers for Realization of Society 5.0” of the Ministry of Education, Culture, Sports, Science and Technology of Japan.

We also thank the students of the Faculty of Engineering, Shinshu University for their cooperation in data collection and provision.

The research is also being partially conducted as Grant-in-Aid for Scientific Research A (19H01102) and C (21K11826).

## REFERENCES

- [1] Shinji Yamane: “Global Game Jam: a case study in the theory and practice of game development in higher education,” IPSJ SIG Technical Report, vol. 2011-CE-108, no. 5, pp.1-6, 2011. (in Japanese)
- [2] Hirokazu Ozaki, Hiroyuki Tominaga, Toshihiro Hayashi, and Hiroyuki Tarumi: “Programming Exercise for Problem Solving with Board-Game Strategy-A Contest Style and the Management Server-,” IPSJ SIG Technical Reports, vol. 2008, no. 2008-EC-9, pp. 1-8, 2013. (in Japanese)
- [3] Takafumi Noguchi: “Programming Education and its Development of Analysis Method which use the Assignment of Making Games,” IEICE, vol. 104, no. 222, pp. 1-6, 2004. (in Japanese)
- [4] Yudai Sugino, Masaaki Niimura, Kozo Okano, and Shinpei Ogata: “Implementation of Programming Environment based on Cloud IDE with Eclipse Che and Docker,” IEICE, Vol.119, No.467, KBSE2019-57, pp.67-72, 2020. (in Japanese)
- [5] Dirk Merkel: “Docker: Lightweight Linux Containers for Consistent Development and Deployment,” Linux Journal, 2014.
- [6] Youhei Matsumoto, Keisuke Fujiwara, and Takehiko Murakawa: “Construction of an e-learning system for information processing education,” Journal of Japan Society of Information and Knowledge, vol. 27, no. 2, pp. 155-160, 2017. (in Japanese)
- [7] Ryotaro Nakata, Kumi Hasegawa, and Yoichi Seto: “Development of Container based virtual exercise system CyExec related to cyber attack and defense,” 2018 Information Processing Society of Japan, vol.2018, no.1, pp.415-416, 2018. (in Japanese)
- [8] Takeru Shimizu, Naoki Hanakawa, and Hiroyuki Tominaga: “Design and Prototype of Execution Environment using Container Type Virtualization for Programming Exercises,” 2019 Information Processing Society of Japan, vol.2019, no.1, pp.535-534, 2019. (in Japanese)
- [9] Zuhal Okan: “Edutainment: is learning at risk?,” British Journal of Educational Technology, vol. 34, no. 3, pp. 255-264, 2003.
- [10] Toru Fujimoto: “Toward Effective Approaches to Educational Use of Digital Games,” Computer and Education, vol. 31, pp.10-15, 2011.

- [11] Ryotaro Nakata, Kumi Hasegawa, and Yoichi Seto: "Development of Container based virtual exercise system CyExec related to cyber attack and defense," 2018 Information Processing Society of Japan, vol.2018, no.1, pp.415-416, 2018. (in Japanese)
- [12] Takeru Shimizu, Naoki Hanakawa, and Hiroyuki Tom-inaga: "Design and Prototype of Execution Environment using Container Type Virtualization for Programming Exercises," 2019 Information Processing Society of Japan, vol.2019, no.1, pp.535-534, 2019. (in Japanese)
- [13] Zuhal Okan: "Edutainment: is learning at risk?," British Journal of Educational Technology, vol. 34, no. 3, pp. 255-264, 2003.
- [14] Ryo Tanaka, Kazuya Tago: "Building hybrid cloud using a container-based virtualization mechanism," 2015 Information Processing Society of Japan, no. 1, pp. 25-26, 2015. (in Japanese)
- [15] Fawaz Paraiso, Stéphanie Challita, Yahya Al-Dhuraibi, and Philippe Merle: "Model-Driven Management of Docker Containers," IEEE 9th International Conference on Cloud Computing (CLOUD), pp. 718-725, 2016.
- [16] Ryan Chamberlain and Jennifer Schommer: "Using Docker to support reproducible research," 5th International Conference on Mechatronics, Materials, Chemistry and Computer Engineering, pp. 1-4, DOI: <https://dx.doi.org/10.6084/m9.figshare.1101910>, 2014.
- [17] Oscar Pedreira, Félix García, Nieves Brisaboa, and Mario Piattini: "Gamification in software engineering – A systematic mapping," Information and Software Technology 57, pp.157–168, 2015.
- [18] Dávid Tengeri, Ferenc Horváth, et al.: "Negative effects of bytecode instrumentation on Java source code coverage," 2016 IEEE 23rd International Conference on Software Analysis, Evolution, and Reengineering (SANER), Vol.1, pp. 225-235, 2016.
- [19] Elinda Kajo and Megi Tartari: "An Evaluation of Java Code Coverage Testing Tools," The Business Continuity Institute, pp. 72-75, 2012.
- [20] Abhinandan H. Patil: "CodeCover: A Code Coverage Tool for Java Projects," ERCICA Elsevier publications, pp.414-421, 2013.
- [21] Ryosuke Tsutsumi, JiuJun Wei, Kozo Okano, Shinpei Ogata, and Masaaki Niimura: "On an Education System for Software with Edutainment based on Eclipse Che and Docker," 27th JSSST Workshop on Foundation of Software Engineering, pp. 149-150, 2020. (in Japanese)
- [22] cdr/code-server, <https://github.com/cdr/code-server> (accessed on April 14, 2021)
- [23] Akanksha Verma, Amita Khatana, et al.: "A Comparative Study of Black Box Testing and White Box Testing," International Journal of Computer Sciences and Engineering, vol. 5, pp. 301-304, 2017.
- [24] Kamran Ghani and John A. Clark: "Automatic Test Data Generation for Multiple Condition and MCDC Coverage," IEEE Computer Society, pp. 152-157, 2009.
- [25] Vard Antinyan, Jesper Derehag, et al: "Mythical Unit Test Coverage," IEEE Software vol. 35, pp. 73-79, 2018.
- [26] Ali Parsai and Serge Demeyer: "Comparing Mutation Coverage Against Branch Coverage in an Industrial Setting," International Journal on Software Tools for Technology Transfer 22, pp. 368-388, 2020.
- [27] Ryohei Takasawa, Kazunori Sakamoto, Hironori Washizaki, and Yoshiaki Fukazawa: "Suggestion of contribution-based gamified testing tool for education," 2013 Information Processing Society of Japan, no. 1, pp. 427-428, 2013. (in Japanese)
- [28] Kazunori Sakamoto, Hironori Washizaki, and Yoshiaki Fukuzawa: "A Programming Contest System for Programming Beginners," Proceedings of the 30th JSSST Annual Conference, pp. 1-8, 2013. (in Japanese)
- [29] Klaus Brandly: "Are You Ready to 'Moodle'?", Language, Learning and Technology, Vol. 9, No. 2, pp. 16–23, 2005.





# Development of a Rainwater Utilization System using LoRaWAN

Shinji Kitagami and Toshihiro Kasai  
Fukui University of Technology, Japan

**Abstract**– In Akashima, one of the Goto Islands in Nagasaki Prefecture, which is a remote island with no groundwater, rainwater is used as domestic water. In order to improve the quality of rainwater to use, it is necessary to collect a wide variety of data on the amount of rainwater and the usage of water in the area. In this paper, we describe a rainwater utilization system using LoRaWAN, one of LPWA (Low Power Wide Area) introduced in Akashima.

**Keywords:** IoT System, LPWA, LoRaWAN, Data Collection, Rainwater Utilization

## 1 INTRODUCTION

As part of the "Akashima Revitalization Project" [1] started in 2017, we are developing a water supply system using rainwater that can be safely used in terms of both water quality and quantity in Akashima, Nagasaki Prefecture. So far, we have constructed a rainwater collection surface, named "Amehata", large rainwater storage tanks, and a pipeline between them.

In order to improve the quality of rainwater used in this system, it is necessary to collect a wide variety of data on the amount of rainwater and the usage of water in the area. For example, to remove the initial rainwater contaminated, the amount of rainwater in storage tank is determined based on analysis results on supply and demand of rainwater. Therefore, it is indispensable to install rainfall sensors in the area and water level sensors at the rainwater storage tanks, and to introduce a mechanism for collecting data related to the use of rainwater. However, in Akashima, there are places where it is difficult to send and receive data via the mobile network. In addition, it is difficult to supply sufficient electric power to sensors at Amehata on the top of the hill.



[Fig.1] Akashima, Goto City, Nagasaki Prefecture

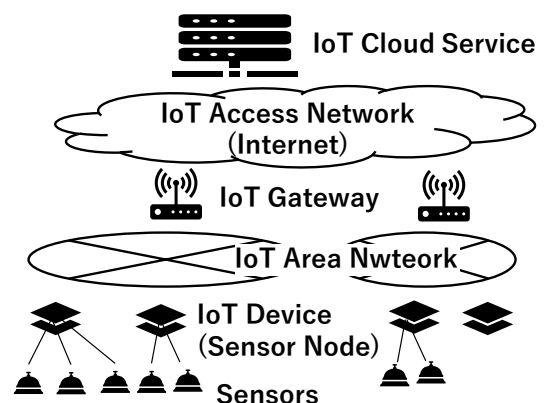
In this study, we have developed a data collection network system to solve above problems. Specifically, we adopted a low power communication technology, LoRaWAN, so that various data can be collected even in places where the mobile network cannot reach. Further, we developed a data aggregation optimization method that collects data with the minimum number of times and size required to further reduce power consumption for data collection [3]. In this paper, we describe the data collection network system introduced in Akashima.

## 2 RELATED TECHNOLOGIES

### 2.1 IoT System

Recently, IoT (Internet of Things) systems have been introduced in various fields, such as industrial and social fields [2]. The IoT system consists of IoT device, IoT gateway, and IoT cloud service as shown in Fig.2. The IoT device connects sensors that measure environmental data such as temperature, humidity and water level to the network. The IoT gateway aggregates many IoT devices. The collected data is visualized and is analyzed by the IoT cloud service. The network between the IoT device and the IoT gateway is called IoT area network, and the network between the IoT gateway and the IoT cloud service is called IoT access network.

Generally, as the IoT area network in factory or home, Wi-Fi or BLE (Bluetooth Low Energy) is used. On the other hand, as IoT access networks mobile phone networks such as 3G / 4G and wired lines such as FTTH (Fiber to The Home) is used. Depending on the IoT system, the IoT device may be directly connected to the IoT access network without installing the IoT gateway.



[Fig.2] IoT System

## 2.2 Wireless Technology for IoT Network

As the IoT area and access network, various wireless technologies are used. These technologies can be classified as shown in Fig. 3 according to the communication speed and reach. Wireless technologies for the mobile phone network such as 3G and 4G (LTE) are often used in outdoor IoT systems because of their high communication speed and long reach. However, as mentioned in Chapter 1, there are many places in Akashima that are out of the service area of the 3G/4G mobile phone network. In addition, as the hardware module for 3G / 4G communication consumes a large amount of electric power, it is difficult to introduce them on the IoT device that cannot supply sufficient power.

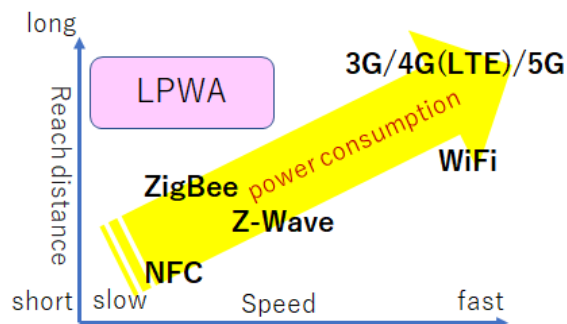
In recent years, LPWA (Low Power Wide Area) has been attracting attention. Using LPWA, the amount of data that can be transmitted per unit time is small, but the reach is long. Also, it is said that one button battery can operate for more than a year that depends on the number of data transmissions.

LPWA is a general name for low power communication methods. Specifically, there are communication standards such as LoRa, SIGFOX and NB-IoT as shown in Fig. 4. Among these standards, NB-IoT requires a license, and mobile phone carriers provide services in general. Therefore, it is difficult to use it in Akashima as well as the 3G / 4G mobile phone network. Also, LoRa is capable of two-way communication, while SIGFOX is capable of only one-way. In our system in Akashima, in consideration of not only data collection but also remote operation of the solenoid valve of the rainwater storage tank in the future, we have adopted LoRa with excellent interactivity.

## 3 DATA COLLECTION NETWORK SYSTEM

### 3.1 System Overview

Figure 5 shows the overview of the data collection network system in Akashima. We installed IoT devices (LoRa sensor nodes), which connected with water flow sensors and water



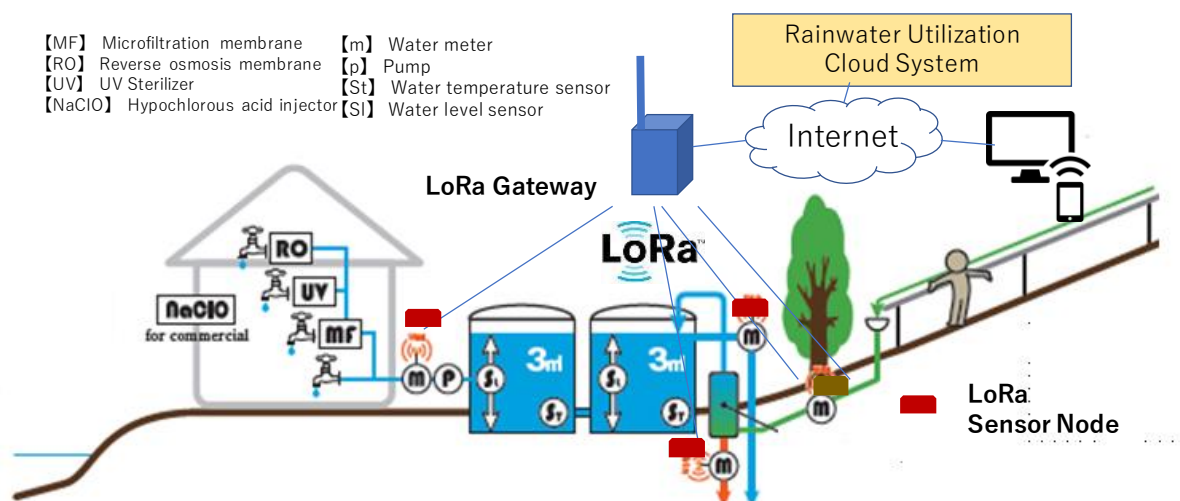
[Fig.3] Positioning of LPWA

Standard	LoRa	SIGFOX	NB-IoT
Reach	11Km	30Km	15Km
Speed	10kbps	0.1kbps	100kbps
License	unnecessary	unnecessary	necessary
Bidirectional	○	×	○

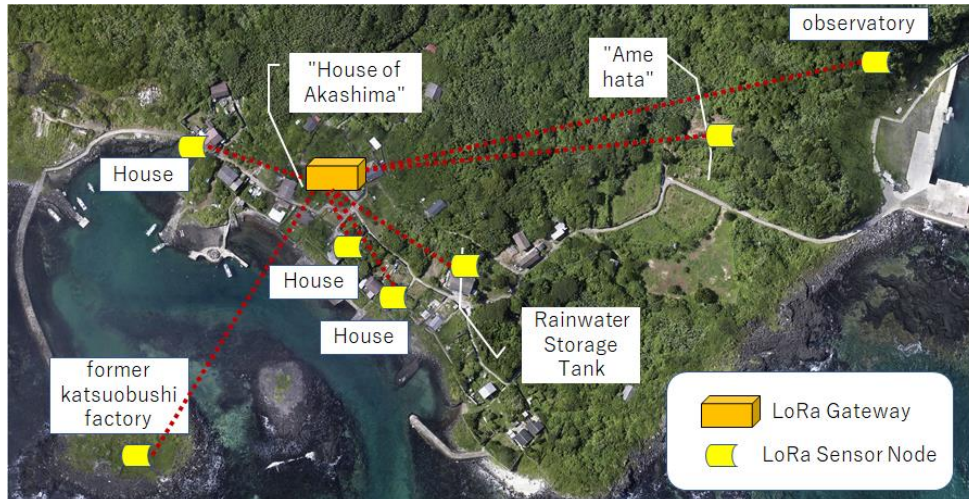
[Fig.4] LPWA standards

level sensors at the rainwater collection surface, named "Amehata", rainwater storage tanks, and private houses that use rainwater. Also, we installed a IoT gateway (LoRa gateway) for aggregating the data measured by the LoRa sensor nodes at "House of Akashima", an accommodation facility for visitors. Communication is performed between the LoRa sensor nodes and the LoRa gateway using the LoRaWAN protocol. The data aggregated in the LoRa gateway is stored in the IoT cloud service platform via the 3G mobile phone network.

Figure 6 shows the locations where the IoT sensor nodes and the LoRa gateway are installed. In addition to the rainwater supply system, we are considering installing LoRa sensor nodes with a rainfall sensor connected to the observatory and the former site of *katsuobushi* factory.



[Fig.5] Data Collection Network System in Akashima



[Fig.6] Installation location of IoT gateway and IoT device

### 3.2 LoRa Gateway

Figure 7 shows the appearance of the LoRa gateway. As hardware platform for LoRa gateway, we adopted the Raspberry Pi 3 Model B +, which is a small board computer, and added an expansion board for LoRa and dongle for 3G communication. DIRAGINO LoRa mini-JP [5] is mounted on the LoRa expansion board, which has acquired technical suitability in Japan. On the other hand, we used SORACOM Air [6] for Cellular for 3G communication.

The following functions are implemented in the LoRa gateway.

- Communication protocol conversion function between IoT area network (LoRa) and IoT access network (TCP/IP)
- Aggregation control function for collected data that minimizes the number and length of communications
- Transmission priority control function according to data characteristics
- Remote maintenance function by on-demand remote access

Here, the aggregation control function of collected data, which is characteristic of the LoRa gateway, is explained as bellow. Conventional data collection methods include a sequential data collection method that collects measurement data each time at regular intervals, a temporary storage data collection method that collects measurement data collectively, and an aggregated data collection method that collects only the aggregated values of measurement data. However, the conventional data collection method has a problem that the power saving property of LoRa cannot be effectively used due to the large size of transmitted data or the frequent transmission [4].

Therefore, we adopted a data aggregation optimization method that autonomously adjusts the data collection interval and the amount of data cache according to changes in the collected data. This method monitors changes in

measurement data at the LoRa gateway and LoRa sensor nodes and sends and receives measurement data only when it matches the data changes defined in advance as conditions. The conditions for sending and receiving measurement data were the following three types.

- (1) **Threshold condition:** Data is transmitted when the data exceeds a preset threshold. When the upper and lower limits of appropriate data are clear, the frequency of data transmission can be minimized.
- (2) **Outlier condition:** Data is transmitted when there is an outlier within a certain interval. It is possible to detect extreme changes in measured values and values that deviate significantly from statistical values.



[Fig.7] LoRa Gateway



- (3) **Rate of change condition:** The value of the measurement data is compared with the value immediately before, and the data is transmitted when the absolute value of the difference exceeds the threshold value. It is possible to detect sudden changes in numerical values, which is effective when the data fluctuates sharply in a short period of time.

### 3.3 LoRa Sensor Node

Figure 8 shows the appearance of the IoT device prototyped in this development. This IoT device is equipped with analog interface, digital interface, and serial communication I/F (I2C) for connecting various sensors. The LoRa communication module DIRAGINO LoRa mini-JP is equipped with ATmega328P as a sensing processing MCU [5].

When the power characteristics of this LoRa sensor node were measured, it was 20mA during data transmission and 12mA during standby. This means that if data is transmitted at 10-minute intervals, it will theoretically operate for about 200 days with three AA batteries. In addition, the operating time can be further extended by using the sleep mode of the MCU. However, in actual use, it is necessary to consider the power consumption of the connected sensor.

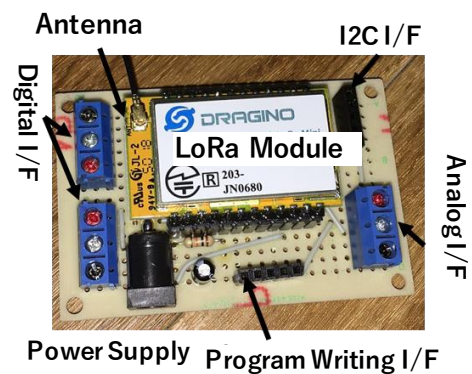
### 3.4 Rainwater Utilization Cloud System

The sensing data measured by the LoRa sensor nodes is transmitted to the rainwater utilization cloud service via the LoRa gateway, and is analyzed and visualized. We use Sakura INTERNET VPS (Virtual Private Server) [7] as the platform of rainwater utilization cloud service.

## 4 FUTURE WORKS

In the future, we plan to install the developed data collection network system in Akashima and evaluate its effectiveness. The future issues that are expected at this point are as follows.

- As for the LoRa specifications, the LoRa gateway and LoRa sensor nodes can communicate even if they are separated by 10 km or more in a good view. However, in Akashima, there is a possibility that radio waves will not reach between the LoRa gateway installed at "House of Akashima" and the LoRa sensor nodes at "Amehata" and the observatory because it is a forest area. In that case, it is necessary to install a repeater for LoRa communication at each intermediate point.
- As a result of preliminary evaluation, the prototyped LoRa sensor node could be operated for about 200 days with three AA batteries. However, to operate the system in an unmanned environment, long-term operability of the LoRa sensor node is required, and it is necessary to install the solar battery.



[Fig.8] LoRa Sensor Node

## 5 CONCLUSION

In this paper, we described the configuration and functions of the data collection network system to be introduced in the rainwater utilization system in Akashima, Goto City. In the future, we plan to confirm its effectiveness and identify future issues.

## ACKNOWLEDGEMENT

Part of this work was carried out under the Cooperative Research Project Program of the Research Institute of Electrical Communication, Tohoku University.

## REFERENCES

- [1] T. Kasai and S. Kondo, "Remote Island promotion project in Akashima, Goto City, Nagasaki Prefecture 2019," Proceedings of the 28th Annual Congress of JRCSA, pp37-41, 2019.
- [2] T. Suganuma, T. Oide, S. Kitagami, K. Sugawara, and N. Shiratori, "Multiagent-Based Flexible Edge Computing Architecture for IoT," IEEE Network, Vol.21, Issue.1, pp. 16-23, Jan. 2018.
- [3] S. Kitagami and T. Kasai, "Development Data Collection System for Rainwater Utilization," the 29th Annual Congress of JRCSA pp14-17, 2020.
- [4] A. Iida, T. Kasai and S. Kitagami, "IoT Data Collection Optimize Method for Rainwater Utilization," Proceedings of the Institute of Electrical Engineers of Japan National Convention, 2019.
- [5] DIRAGINO LoRa mini-JP, <http://www.openwave.co.jp/lorawan/>.
- [6] SORACOM Air for Cellular, <https://soracom.jp/services/air/cellular/>.
- [7] Sakura INTERNET VPS, <https://vps.sakura.ad.jp/>

# Design- thinking information system development methods for the information-impaired people

Koji Yamada\*, Toru Takahashi\*, Sakiko Kasuya\*\*, and Nobuhiro Kataoka\*\*\*

\*Faculty of Design Technology, Osaka Sangyo University, Japan

\*\*Gifu Shotoku Gakuen University Junior College, Japan

\*\*\*Enterprise Laboratory, Japan

{yamada, takahashi}@ise.osaka-sandai.ac.jp

sakiko@gifu.shotoku.ac.jp

kataoka9@kataoka9.com

**Abstract** - There are a certain number of information-impaired people who cannot take the plunge into information technology, despite the fact that information technology can improve their business. The information-impaired people in this paper are those who cannot benefit from information technology due to lack of personnel or budget in the organization. In order to solve this problem, we developed an information system development method based on design thinking. We propose a method in which the information-impaired people can develop, implement, and operate information systems themselves. We focused on the teachers and staff of child-care facilities as one of the categories of the informationally vulnerable, applied the proposed method to their work, and confirmed the effectiveness of the method. We applied the proposed method to the pick-up management of a childcare facility and started trial production in January 2020. As a result, in March 2021, we obtained trust in the information system and suggestions for improving its functions from the information-impaired people.

**Keywords:** Design thinking. Information-impaired people. Information system development method. Child-care facility. Improvement of business

## 1 INTRODUCTION

There are a certain number of information-impaired people who are unable to take the initiative in using information systems, despite the fact that information technology can improve their business. The reasons for this are the lack of funds in their organizations and the lack of support for information technology. Also, there is a lack of skills among the information-impaired people themselves. We focused on teachers and staff at preschools, kindergartens, nursery schools, and other early childhood facilities as one category of information-impaired people, and in 2018, we published a report on the lack of information technology in early childhood facilities [1]. Since then, with the expansion of subsidies and other programs, packaged software to improve the operations of childcare facilities has become more popular [2][3].

However, teachers and staff of child-care facilities do not have the professional skills to evaluate the effectiveness of packaged software. Nor do they have the time to do so. There is also packaged software that can be customized to

meet the needs of each child-care facility, but this comes at a cost. In some cases, customization is not possible and a proprietary information system must be implemented. As a result, the teachers and staff of childcare facilities, who are information-impaired people, give up on using the information system. They lose motivation to improve their work.

To solve this problem, we propose a method based on design thinking that allows information-impaired people to develop their own information systems. In this way, information-impaired people can implement their own information systems. By using this system, they can improve their business. We applied the proposed method to the business of a child-care facility and showed its effectiveness in an empirical experiment.

Chapter 2 describes the proposed method, Chapter 3 describes the application of the proposed method, Chapter 4 describes the empirical experiment, Chapter 5 describes the evaluation of the proposed method, Chapter 6 describes the results, and finally, the summary is given.

## 2 PROPOSE A SOLUTION TO THE PROBLEM

### 2.1 What is Design Thinking?

Design is not an extension of "industrial design" of the 20th century industrial society, but should be called "knowledge design" of the 21st century knowledge society. Design thinking is a process that encourages intuitive and comprehensive human-centered thinking in order to construct knowledge design [4]. The process of design thinking begins by intuitively hypothesizing some episodes related to the subject (e.g., a workplace that is constantly working overtime, long meetings, etc.). Then, through involvement with related surrounding objects and information, and interaction with people, the episodes are revised to form a certain understanding. In doing so, we often go back to the starting point. Eventually, the elements and knowledge surrounding the subject are organized in an implicit way.

In the existing research on design thinking, after an overview of theories and discussions, the forms and issues of design thinking that are currently required are described

[4]. There are also reports of practice and verification based on design thinking [5]. On the other hand, there are some reports on the application of design thinking to system development, but there are no reports on detailed procedures for information system development [6].

## 2.2 Development methods by information-impaired people themselves

In order to conceptualize the method of system development by information-impaired people, we referred to the book "Innovate by Design-based Management" [7]. In this book, the process of design thinking is explained using an observational engineering approach: observation, conceptualization (hypothesis), and prototyping. It states that hypotheses are developed through descriptive and verbal research such as fieldwork, interviews, and case studies (Fig. 1).

On the other hand, in the reference, the methodology is described in detail, but the results of the design thinking process are described categorically only as a result of applying the methodology, without discussing empirical examples. In order to apply the Design Thinking Process, we thought it was important to present a more concrete methodology. In addition, in the conventional information system development method, whether it is the waterfall model or the agile model, the requirements are defined in detail. At this time, optimization of the information system is pursued in order to achieve the implementation effect, such as redefining the business process. The system is designed in an extremely comprehensive and decisive way, including a close examination of the documentation of the current system and interviews with management and employees. In contrast, in the design thinking process, an intuitive hypothesis is formulated based on several episodes. Then, observations are made and conceptualization (concept formation) is carried out through intuitive thinking. Then, necessary and effective functions are implemented and operated through repeated prototyping. We believe that this method will lower the hurdle for the information-impaired people to develop their own information systems and position information system development as one of their jobs.

Figure 2 shows a proposal for a method of information system development by information-impaired people

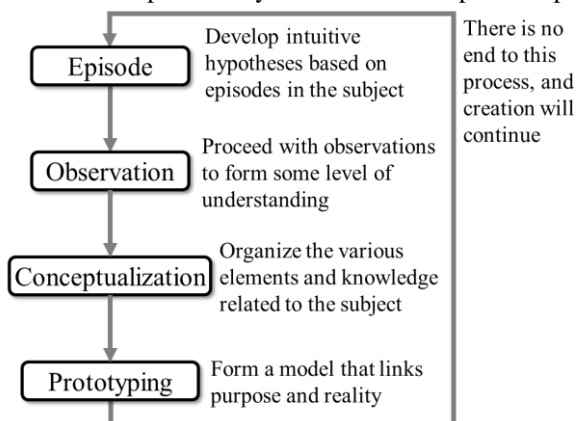


Figure 1: Design Thinking Process

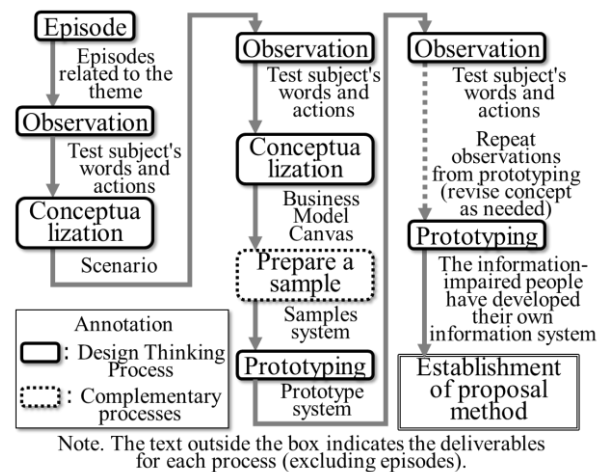


Figure 2: Information system development method by information-impaired people themselves

themselves, based on the design thinking process shown in Figure 1. It is the same as the previous method in that the information system development experts participate in the process at first, but the role of leading the process is shifted to the information-impaired people. Finally, it is a way to realize the development of information systems in which the information-impaired people themselves contribute to their own business. In this paper, we have defined the artifacts of an experiment to demonstrate this proposal. A scenario is created as a result of observation and conceptualization based on episodes. Further observations are made, and the concept is developed again. This artifact is the business model canvas. In order to realize the proposed method, there are people and events involved, mainly information-impaired people. These will be defined. This process is called "prototyping" in the literature, and we believe that in order for information-impaired people to perform prototyping, we need "things" to support the formation of the image. Therefore, we created a sample system, applied it to prototyping, and observed the behavior of the information system. This process of prototyping and observation was repeated. Eventually, we will lead the users, who are information-impaired people, to a state where they can add functions to the information system by themselves with advice from experts in information system development, and also to a state where they can develop the information system by themselves.

### 3 APPLICATION OF THE PROPOSED METHOD

### 3.1 Episode, observation, and Conceptualization

The episode we focused on was the following experience of Kasuya, one of the authors. Mrs. Kasuya's main job is to teach at a junior college that trains teaching staff for child-care facilities. In order to secure practical training for her students, she had to interact with child-care facilities. On the other hand, the use of IT at child-care facilities has not progressed, and she has witnessed the exhaustion of teachers



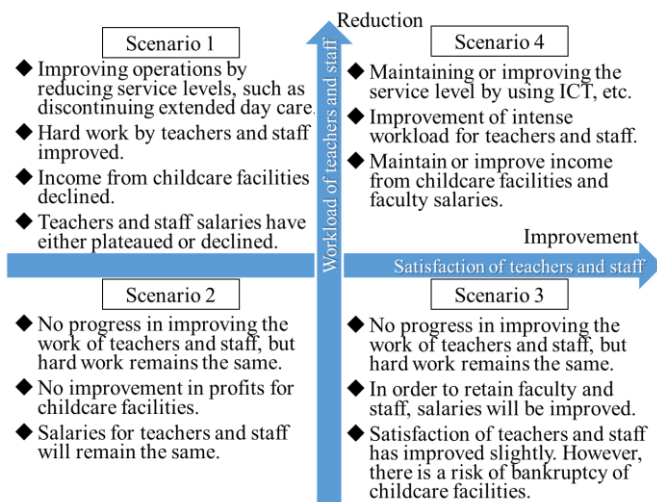


Figure 3: Results of scenario planning

and staff due to the many manual tasks involved in preparing for events such as fields days and answering phone calls from parents to confirm their children's attendance.

In October 2017, as an inspection of childcare facilities in Gifu Prefecture, we conducted a survey on the current status of information technology in childcare facilities. First, as a preliminary survey, we conducted a questionnaire for participants of a course for faculty members conducted at a university. In addition, we interviewed two child-care facilities that have introduced commercially available software packages. We then conducted this survey by sending paper questionnaires to 560 childcare facilities in Gifu Prefecture. We received responses from 248 facilities (44%). 77 facilities (31%) used packaged software. We also found that spreadsheet software and word processors, which are also used by general companies, are being used for informatization in childcare facilities. On the other hand, there was no progress in the computerization of tasks characteristic of childcare facilities, such as attendance management and communication with parents. The reason for this was, in a word, anxiety and denial due to "not knowing". Lack of clarity about the benefits and costs of digitization. Security concerns and PC skill concerns [1]. Based on the episodes described above, we defined the information-impaired people as the subjects in this study as teachers and staff (hereinafter referred to as "teachers and staff") of childcare facilities.

Then, we developed the concept. Using a scenario-based approach, we drew several scenarios. The scenario approach is to design events that happen in time and space, or "possible" events. Rather than predicting the future through analysis or narrowing down "certain" elements, the goal is to input as many possibilities and uncertainties as possible to add depth to strategies and concepts, and to draw out areas of innovation and possibilities for business development [7]. The method we used for this purpose is called scenario planning. This is a method of discovering uncertain factors that will affect the future, representing the future in multiple scenarios, and envisioning actions in advance.

As a preliminary step to scenario planning, we conducted a survey of existing public documents and literature [8].

Based on this survey, we set two axes for scenario planning: "Satisfaction of teachers and staff (improvement/decline)" and "Workload of teachers and staff (reduction/maintenance/worsening)". For each quadrant of these two axes, we drew the relative profitability of the childcare facility and the salary level of the teachers and staff (Figure 3). As a result, we came to the conclusion that we should adopt scenario 4, which is to increase satisfaction of teachers and staff and reduce their workload.

### 3.2 Observation (2nd time)

Following the results of the scenario planning, we conducted a survey on the current situation of childcare facilities. The purpose of the survey was to get their opinions on the scenarios and to obtain information that would lead to additions and modifications to the scenarios and to the design of business models and prototypes.

Survey 1: The current status of the management of childcare facilities that students experience in their training

(1) Survey target: Five second-year students in the Department of Early Childhood Education at Gifu Seitoku Gakuen Junior College

(2) Survey date: August 4, 2018 (Saturday)

(3) Survey method: Face-to-face survey of the current status, issues, and suggestions for improvement in the management of childcare facilities experienced by students during their childcare practice.

(4) Results: The students appealed for improvement in their work through information technology during their childcare practice. However, no improvement was granted. They were opposed by old teachers and staff who had an aversion to information tools. Their reasons were extremely emotional: that they could not feel the warmth unless they wrote by hand, and that using information tools was cutting corners.

Survey 2: Listening to the opinions of government officials involved in the operation of childcare facilities on the results of scenario planning (expecting the introduction of childcare facilities that can be prototyped)

(1) Survey target: Two persons in charge of early childhood policy in Y City, Osaka Prefecture

(2) Survey date: August 14, 2018 (Tuesday)

(3) Survey method: We explained the scenario and collected free opinions from the perspective of guiding the operation of child-care facilities as a government agency.

(4) Results: The discussion about the scenario was not deepened. The reason for this was that we were busy preparing for the introduction of the preschool management system for the public childcare facilities to be developed under the leadership of Y City in the next fiscal year, and could not take enough time. For the same reason, we were also unable to get an introduction to a childcare facility where prototyping could be done.

The information gleaned from these surveys is that teachers and staff unconditionally accept their existing jobs and are strongly resistant to improvements in their work. Attendance and attendance management is a task performed at many childcare facilities, and there are high expectations

for improvement. On the other hand, government-led business improvement inevitably leads to uniformity, and business improvement led by childcare facilities is difficult.

Public child-care facilities are managed by the government, and it is not possible to promote IT at specific facilities alone. In addition, it is difficult to keep track of income and expenditures. In other words, the effect of the proposed method on the management of childcare facilities cannot be measured. Private child-care facilities, on the other hand, can make their own efforts. In addition, the income from parents, subsidies, and expenses are clearly defined. Therefore, we decided to focus on private childcare facilities in this study.

### 3.3 Conceptualization (2nd time)

Following the results of the scenario planning, we need to link the results of this study to the value of the subjects. That is to improve the satisfaction of the teachers and staff and to reduce the workload of the teachers and staff. On the other hand, it is also important that the proposed method can be applied to many childcare facilities. Therefore, we decided to examine the business model from the reference literature. A business model is essentially a design of the relationships among various elements that link the company's activities to customer value [7]. Therefore, in this study, we created a business model canvas in order to increase the satisfaction of the teachers and staff and reduce their workload, and considered the relationships among various elements to achieve this. The results are shown in Figure 4.

The first element is "Products & Services". Customer value is important for this. We defined this as "a business support system for private childcare facilities that is planned, designed, manufactured, and operated primarily by teachers and staff" ("Childcare facility business support system"). information-impaired people find it difficult to write code in C, JAVA, and other languages. In recent years, however, tools that enable the construction and operation of information systems with less code or without writing code have been put to practical use. Therefore, we propose a low-code/no-code development tool for the core of products and services. In this proposed methodology, we have adopted Cybozu's kintone[9]. We chose to use kintone[9] by Cybozu, Inc. because kintone was the pioneer of this tool and was used by nearly 10,000 companies at the time, and also

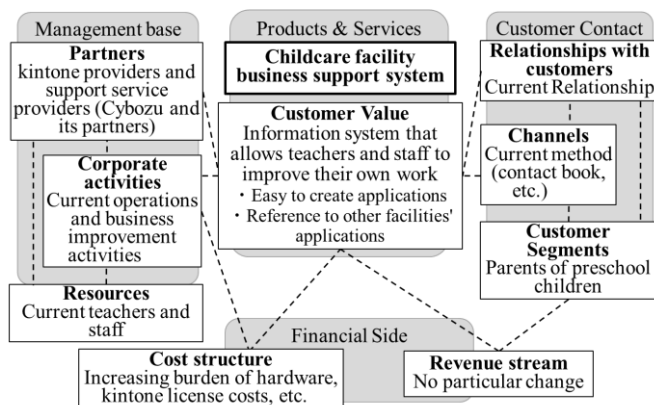


Figure 4: Business Model Canvas for Proposal Method

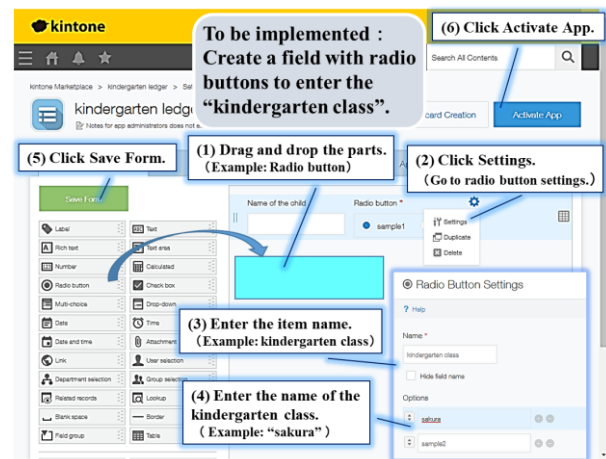


Figure 5: Outline the steps to create an application using kintone

because the kintone website has extensive support content. Building a system, called an "application," can be accomplished by simply dragging and dropping text input, radio buttons, and other parts into a form on a browser.

It shows how a person developing, for example, a kindergarten ledger consisting of elements such as name, address, and parent information. After logging in and instructing application creation from the portal screen, the screen shown in Figure 5 is accessed. The developer moves the elements of the child ledger from the parts list on the left of the screen to the form on the right of the screen by dragging and dropping. Next, enter the contents of the elements and the items to be selected in text format like a word processor (2) to (4). Repeat these steps. Finally, click Save Form (5). Continue by clicking (6) Activate App. With these extremely simple operations, an information system can be created.

In the "management base" section, we designed corporate activities, resources, and partners. For corporate activities, we defined business improvement activities in addition to the current operation of the childcare facility. Resources are the teachers and staff who work at the childcare facility. Partners are companies that provide advice to help improve the operations of childcare facilities and teachers. kintone has many Cybozu partners that provide this function. In addition, we have established a Cybozu developer network for "knowledge transfer" via the Internet.

Customer contact" is designed in terms of customer segments, channels, and relationships with customers. A customer segment is a parent of a preschool child to whom customer value is provided. Channels are the media and communication channels through which customers are reached and value is provided. This is basically the current method (e.g., contact books), but it is also subject to business improvement. Relationships with customers are also traditional, but to improve the work of teachers and staff, parents of children who are customers may also become users of the information system.

On the "financial side" are the revenue stream and cost structure. There is no significant change in the revenue stream. Although there is no significant change in the revenue stream and the cost structure, we need to consider the cost of implementing the information system by

applying the proposed method. This includes the hardware to operate the information system and the license for kintone. In addition, WiFi facilities may need to be installed.

## 4 PROTOTYPING

In order to clarify the business issues specific to the child-care facilities targeted for prototyping, we thought that a sample system would be necessary. Therefore, we have created a sample system. We have created four types of systems: (1) an attendance management system, (2) a system that provides information on the location of buses used for attendance, (3) a system that measures and visualizes the physical activity of the staff of child-care facilities to equalize their workload, and (4) childcare records, which are generally provided as packaged software [11].

The reasons for using these as sample systems are as follows: 1) An attendance management system is an essential task for nursery school among child-care facilities. Unlike kindergartens, the childcare hours are different for each child. 2), 3) This is a task that we believe has the potential to improve the operations of child-care facilities by applying the technology we developed in another research project. 4) Childcare records are one of the pieces of information that are linked when a child graduates from preschool and enters elementary school. Therefore, it is a task that occurs in any childcare facility.

From November 2018, Kasuya took the lead in recruiting subjects for the demonstration experiment to childcare facilities in Gifu Prefecture.

### 4.1 Management of childcare and drop-off at K kindergarten

We made our first visit to K kindergarten in June 2019. The kindergarten is open from 9:00. to 14:30. Children arrive at the kindergarten in one of several buses scheduled to arrive at 9:00. The bus takes the children to the kindergarten. The children will then participate in kindergarten classes such as drawing, music, English conversation, etc., before getting on the bus again at 14:30

for drop-off. There is no need to keep track of the children's arrival and departure times, except in the case of absences.

On the other hand, there is a service called "extended childcare". In consideration of the recent situation of working parents, kindergartens will take care of children after 14:30. In other words, the kindergarten replaces the function of the nursery school. In this case, parents must go to the kindergarten to pick up their children. The time they pick up their children must be recorded in minutes. The reason for this is that the fees for extended childcare vary depending on the drop-off time. At that time, the parents who came to pick up their children pressed the time cards installed in front of the staff room. This data was tabulated monthly for each child. In addition, we calculated the fees for extended childcare. In the prototype, we first introduced and operated a system to ensure the recording of drop-offs and pick-ups, and after the system was established, we decided to streamline the calculation of extended childcare fees.

In this way, we surveyed the K-kindergarten and reached a consensus on the details of implementation. Through this process, we decided on a specific way to proceed with the demonstration of the information system development method by the information-impaired people themselves, as outlined in Figure 6. After the third observation in Figure 2, We simply described the prototyping process as repeating from observation to prototyping. The dashed line in Figure 6 embodies what was described as simply repeating. We will ask the teachers and staff for their opinions on further improvements, and revise the information system to contribute to the improvements. By repeating this process, the teachers and staff will improve their own awareness of the need to improve their own work, and raise their awareness of the need to start developing their own information systems.

We have identified specific issues that need to be improved by November 2019. Initially, we thought that we only needed to obtain records of drop-off times, but there were two ancillary tasks involved in managing childcare drop-offs: one was to open the gates of the kindergarten, and the other was to ensure that the kindergarten's doors were open. Strict security is required for childcare facilities. No one

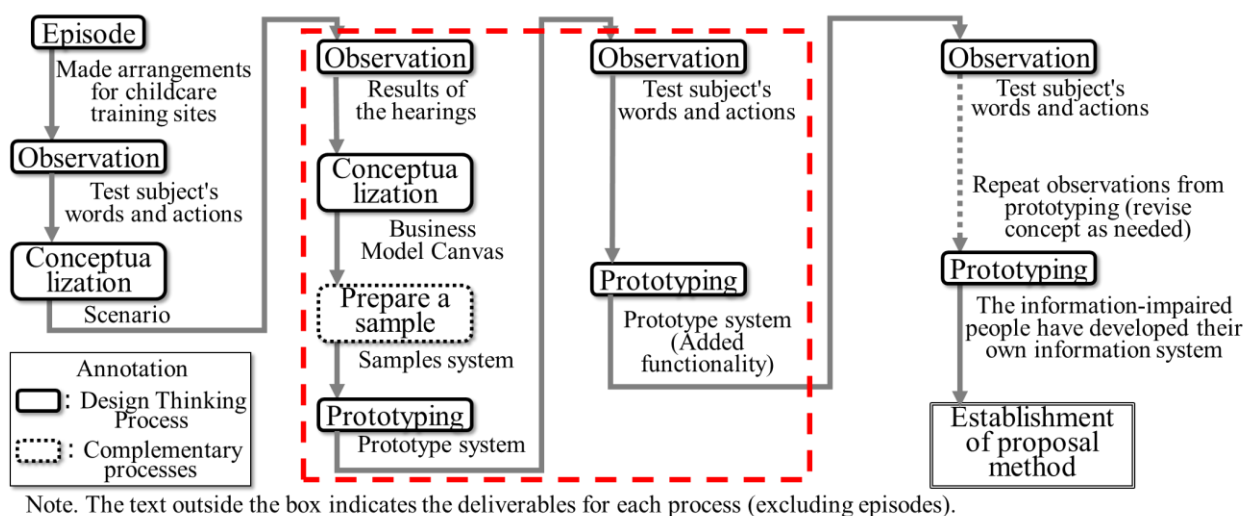


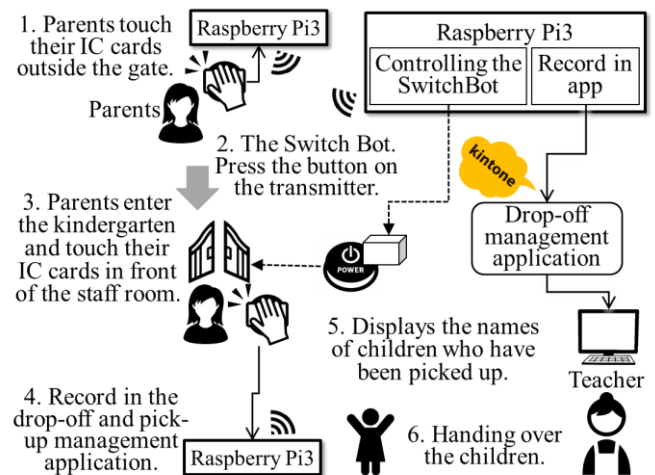
Figure 6: By the teachers and staff of childcare facilities themselves Information system development method



other than parents are allowed to enter the kindergarten, and the gates are locked. When parents go to pick up their children from the kindergarten, they must use the intercom outside the gate to notify the teachers and staff of their child's arrival. Teachers and staff stop what they are doing and answer the call. Using the low-power wireless communication system installed in the existing gate to open the electric lock, the gate opens when the button that sends the opening signal is pressed. This was a task that occurred about 60 times a day, from about 15:30 to 19:00.

In the second case, teachers and staff call children who have been picked up by their parents at the kindergarten. The children in the kindergarten are in a classroom away from the staff room. Therefore, the teachers and staff who just answered the phone will tell the teachers and staff in the extended childcare classroom who the guardian is via the campus phone. At this time, the teachers and staff of the extended childcare classroom were often unable to answer the phone because they were busy dealing with the children. In some cases, parents would call on the intercom outside the gate, open the door, and enter the kindergarten with the parents who were right behind them. At this time, the teachers and staff would know for the first time which preschooler's parents they were in the classroom. From there, they would prepare to leave, and it took a long time to get the preschoolers off the premises.

In order to improve these tasks, we have created a prototype system that links the Raspberry Pi to kintone, connects an IC card reader to the Raspberry Pi, reads the data from the IC card, and uses it as a record for the kintone application. For the former task, when the IC card is touched to the reader at the gate, the Switch Bot [10] is used to push the button of the lock-opening signal transmitter. For the latter task, when the IC card is touched to the reader in front of the staff room, the name of the child whose parent or guardian has come to pick up the child is displayed on the display in the childcare classroom, and the child is notified



Note. The Raspberry Pi3s communicate with each other via WiFi.

Figure 7: Process for drop-off of extended childcare

of which child has been picked up [11] (Figure 7).

In January 2020, we conducted a trial operation for about 10 parents, and in early February, we explained to the parents about the introduction and operation of the system. the system was opened to the public on February 18, and has been in operation for one year and three months without any major problems.

## 4.2 Added a function to calculate extended childcare fees

From June 2020, we have improved the method of calculating the extended childcare fee. In the current system, teachers and staff visually extract the drop-off time from the time cards and calculate the extended childcare fee. The extended childcare fee is then calculated based on the extended childcare fee regulations and entered into an excel sheet. By downloading the data from the kindergarten drop-

Table 1: Evaluating each element of the business model canvas

Area	Element	Contents	Current situation	Evaluation
Products & Services	Customer Value	Information system that allows f teachers and staff to improve their own work • Easy to create applications • Reference to other facilities' applications	Prototype system has been built and is being used in two kindergartens for a demonstration experiment. Not yet built by teachers and staff.	Can't say it's good or bad
	Partners	kintone providers and support service providers (Cybozu and its partners)	Established relationships with Cybozu and its partners	Good
Management base	Resources	Current teachers and staff	No change	Fairly good
	Corporate activities	Current operations and business improvement activities	Raising awareness of business improvement	Fairly good
	Relationships with customers	Current Relationship	The author has started to communicate with them through his main business.	Fairly good
Customer Contact	Customer Segments	Parents of preschool children	Teachers, staff and parents of preschool children	Fairly good
	Channels	Current method (contact book, etc.)	Report on research progress via SNS and web media	Good
	Cost structure	Increasing burden of hardware, kintone license costs, etc.	Not yet considered	Not good
Financial Side	Revenue stream	No particular change	Not yet considered	Not good

off and pick-up management application and posting it to the Excel sheet, we were able to provide and operate the Excel sheet for calculating the extended childcare fee [11].

Teachers and staff learned how to work by referring to tutorials; on October 1, 2020, we experimented with creating a list of fees for the previous month. As a result, this task was successful. Also, since this task occurs once a month, we needed to establish the skills to perform the task. After that, we continued the experiment, and although there were temporary inquiries to us in November and December 2020, we were able to solve them by ourselves and successfully carry out the work.

## 5 EVALUATION

The prototype system has been in operation for one year and three months without any major problems; in March 2021, the system was in operation for 21 days; the number of lock openings and drop-offs per day was 40 and 51, respectively. Next, we assessed the current status of each of the nine elements of the four domains defined in the business model canvas (Table 1).

The construction of a new application by the teachers and staff themselves, has not yet been realized. On the other hand, we have received new requests from teachers for the prototype system. Specifically, the names of the children who have been picked up are displayed on the display in the childcare classroom, but the children sometimes overlook them. In order to improve this situation, they requested that we install a patrol light in the blind spot of the display to notify them. This was not a request obtained from the hearing, but was raised spontaneously by the teachers and staff. This can be said that the teachers accepted the use of the prototype system. Therefore, we rate the product & service as "Can't say it's good or bad". As for the management base, we succeeded in communicating to the president and other members of the management team of Cybozu that the system created by the proposal method was undergoing demonstration tests. At the same time, we are strengthening our relationship with our partners and building an infrastructure that will allow us to continue to receive support from them. In terms of customer relations at the point of contact, we rate the prototype system as "good" because it contributes to the convenience of not only teachers and staff but also parents. In terms of financial side, there is no movement at present.

## 6 DISCUSSION

This section discusses the effectiveness of the proposed method.

The purpose of the proposed method is to develop an information system by the information-impaired people themselves through repeated prototyping and observation after the conceptualization. As a result of the initial prototyping, teachers and staff raised a new business problem. In a kindergarten pick-up and drop-off management system, teachers sometimes miss the names of the children who have been picked up from the kindergarten even though they are shown on the display. To improve this situation, a suggestion was made to use a notification

method other than the display. One teacher suggested using a patrol light. We believe that this raised the teacher's own awareness of the need to improve their work and their expectations of the information system. In terms of the phases of information system development, this can be considered the planning and basic design stage. By repeating the trial production and observation, we believe that the possibility of teachers and staff themselves developing the information system has increased.

Next is the expansion of the scope of application of the information system. In March 2021, teachers and staff informed us that they would like to apply the system to "morning" childcare, and from April of the same year, they would like to discontinue the time card system they had been using. We believe that this is another event that shows that the teachers and staff have the desire to solve problems on their own. We also believe that their decision to discontinue the parallel operation of time cards is a sign of their confidence in the ability of the information system to improve their work.

In addition, almost all communication between teachers, staff and us was done through a chat function called kintone thread. The purpose of the proposed method is for the information-impaired people themselves to develop the information system. Being able to communicate in a thread means that advice and specific development methods are more likely to be conveyed. We believe that the use of threads will foster the skills necessary to enable the development of information systems by information-impaired people themselves.

As a result, we have realized the following (Table 2) as steps for system development by the information-impaired people themselves.

Table 2: Evaluation of the proposed method

Evaluation Item	Evaluation
Trust in the information system by the information-impaired people	Good
Operation of the information system by the information-impaired people	Good
Proposal of information system functions by the information-impaired people	Fairly good
Development of information system by the information-impaired people themselves	Not so good

## 7 SUMMARY AND FUTURE PROSPECTS

Based on design thinking, we developed a method for developing information systems that can be improved by information-impaired people themselves. The process of observation, conceptualization, and prototyping was repeated using teachers and staff of a childcare facility as test subjects. As a result, we were able to obtain the opinions of the teachers and staff about improving their work and their intention to expand the scope of application of the information system. As a result, we found that the proposed method may be effective as an information system development method for information-impaired people.

In the future, we would like to meet the needs of the teachers and staff of K kindergarten. For the improvement

of the notification method, the teachers suggested the Patlite. In addition we also suggested pagers and bone conduction earphones. We will then verify if the proposed system is effective in other childcare facilities. We will summarize the results of the prototype system and obtain new subjects. To this end, we began approaching the city of D in Osaka Prefecture. At the same time, we will explore the factors that lead teachers and staff to take the initiative in developing an information system and review the concept.

## REFERENCES

- [1] S.Kasuya, Issues of computerization of school affairs in nursery schools and kindergartens, Electronics, Information and Communication Engineers, IEICE General Conference 2018, D-15-21 (2018).
- [2] CoDMON, Inc. , CoDMON, <https://www.codmon.com>, (Access on May 23,2021).
- [3] UniFa, Inc. , lookmee, <https://lookmee.jp>, (Access on May 23,2021).
- [4] T.Taura, Issues of computerization of school affairs in nursery schools and kindergartens, Transdisciplinary Federation of Science and Technology, Oukan, Vol.12, No.1, pp.5-13(2016).
- [5] K.Onai, N.Mori, and K.Sugiyama, Studies on Modelings of The Thinking Proceses of Design (Part 2), Japanese Society for the Science of Design, Journal of the science of design, Vol.1988, Vol.68, P.43(1988).
- [6] T.Yarimizu, H.Kitanaka, Study on the structural changes "Design Thinking" brings to the IT system development, Japan Society for Management Information, Abstracts of Annual Conference of JASMIN, pp.186-187(2019).
- [7] N.Konno, Innovate by Design-based Management, Toyo Keizai, Inc. (2010).
- [8] K.Yamada, T.Takahashi, and S.Kasuya, Implementation of information system by child care facility operation support system concept and fast system based on design thinking, The Institute of Electronics, Information and Communication Engineers, IEICE technical report, Vol.119, No.66, pp.27-32(2019).
- [9] Cybozu, Inc. ,kintone, <https://kintone.cybozu.co.jp/>, (Access on May 23,2021).
- [10] Wonderlabs Inc. ,SwitchBot Curtain, <https://www.switch-bot.com/>(Access on Dec.23,2020).
- [11] K.Yamada, K.Yamaoka, T.Takahashi, and S.Kasuya, Report on the contents of the demonstration experiment and the results of the initial stage in the nursery facility business support system concept, Electronics, Information and Communication Engineers, IEICE technical report, Vol.120, No.376, pp.8-15(2021).

Session 8:  
Application II  
( Chair: Yoshia Saito )





# A proposal of method for collecting daily physical and mental state data using communication tools among parenting generation

KANAE MATSUI<sup>†\*</sup> KAZUMA NISHIGAKI<sup>†\*\*</sup>

<sup>\*</sup>School of System Design and Technology, Tokyo Denki University, Japan

<sup>\*\*</sup>Graduate School of Informatics, Tokyo Denki University, Japan  
{23jkm23@ms, matsui@mail}.dendai.ac.jp

**Abstract** – In this study, we propose an IoT-based platform to collect data of males and females who are experiencing changes in their life stages. The platform has two functions; one is to collect the data for providing information, which helps to maintain their daily health, the other is to provide the information. The proposed system is contained with wearable and IoT devices, and chat tool, Slack to acquire both physical and mental data. This paper describes the data collection part. We conducted a feasibility study to check its applicability.

**Keywords:** IoT application, Healthcare, Data collection

## 1 INTRODUCTION

Males and females in grades with small children are expected to play a leading role in economic activity. Also, they have many changes in life stages such as marriage, child-rearing, and promotion at work, physical and mental illnesses, such as lifestyle-related diseases and mental depression. Reference [1] [2] describes their risks. It is important to be in good health before suffering a major illness from the viewpoint of non-illness [3], and since the physical and mental conditions change daily, continuous monitoring of the living environment is necessary. A possible monitoring method is the use of devices, systems, and applications that use Internet of Things (IoT) technology.

With the spread of IoT technology, it has become possible to acquire a wide variety of fine-grained data and provide feedback to humans and machines using the collected data [4]. In addition, concepts such as smart cities, communities, and houses that support people's lives using IoT data have emerged, and applications have been implemented. The applications use indoor or outdoor environment data, which obtained from network-compatible sensors. In addition, with the spread of wearable devices and IoT devices, vital data such as pulse, blood pressure, and body temperature that can be collected from the human body and utilize them for medical treatment and healthcare.

In addition, the IoT data is not defined only as obtained from sensor devices, but also includes social networking services (SNS) and web questionnaire data [5]. Since it is difficult to measure and estimate the mental state only from the IoT device, it is necessary to obtain self-evaluation data of the physical and mental health state by using a questionnaire, which is working with the chatbot format or a web questionnaire.

Therefore, the authors collect the physical condition by IoT device, collect the mental condition by web questionnaire, monitor the current physical and mental condition. Our purpose is to verify the possibility of providing information on what types of consciousness change and behavior change will be occurred by provided information.

We aim to develop proposed platform that collects and provides data and explain how the platform is useful for the purpose. The proposed platform has two main functions; one is to collect physical and mental condition data, the other is to provide the information for maintaining or improving the condition. This paper focuses on former part of the data collection, because if there is an error in the collected data, the provided information will not have reliability. Therefore, we first verify functions of the platform in this paper, also Figure 1 shows the position of this paper.

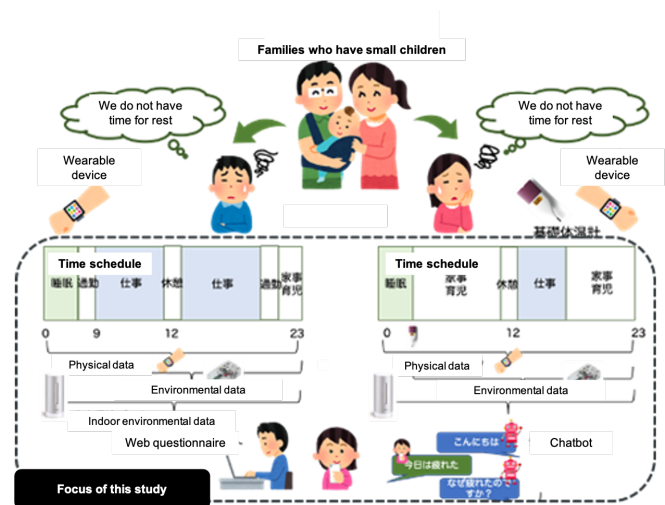


Figure 1 Overview of this study and focus point.

To check the availability of this function, we introduced the 10 people in 5 households to use it for about a month.

## 2 RELATED STUDY

In this chapter, we describe behavioral changes for maintaining health condition using IoT devices and wearable devices and clarify the position of this study.

### 2.1 Maintaining indoor comfort

Reference [6] introduces smart homes uses IoT devices and the data collected using them to maintain indoor comfort. In

smart homes, energy management system (EMS) using smart meters is a main target. However, comfort in household is often prioritized then economic benefits. Therefore, smart home systems that realize energy saving while maintaining indoor comfort have been developed and introduced. However, indoor comfort is diverse, unlike office comfort, which is different from productivity indicators. Therefore, machine-learning has been used for finding each person's comfort zone [7].

## 2.2 Maintaining health using physical data

Technological innovation in wearable devices has made it easier to collect physical data. Reference [8] introduces that not only physical data is used for determining the amount of activity, but also attempted to determine physicality and mental satisfaction, such as a well-being state. For that, data collected by wearable devices and behavior recognition technology are used [9], and data from introspection determined by person such as physical data and mental state data are used.

## 2.3 Position of this study

When IoT-based measurement technology has been established, environment data, such as indoors and outdoors, can be sensed and collected. Therefore, we designed a seamless system that can be adapted to a series of life as sensing targets. Because people's life is not limited to indoors and outdoors, and a series of life flows affect the body and mind. Currently, devices of wearable, IoT, and smartphones can acquire this continuous body status data. By these

methods, we constructed a platform for collecting data related to a series of human behaviors including their surrounded indoor and outdoor environmental data.

Also, for collecting the mental condition, judgment their mental condition by others such as counseling and coaching, or self-judgment such as introspection is required. Therefore, we use a communication chat tool (Slack [10]) and the body of the day, 7 stages of mental state data is collected by selecting from in Slack (Figure 4).

The main purpose of this study is to develop an information providing system for the purpose of maintaining health for generations that undergo major changes in their life stages, such as the child-rearing generation. The system has two types of information presentation, physical and mental state data. Before developing the information provision system, we need a platform which collect both physical and mental data. Therefore, we introduce the platform that collects these two types of data using IoT devices and Slack, and then verify their accuracy of data collection.

## 3 PROPOSED PLATFORM

The proposed platform has a function of collecting the physical data by wearable devices and environmental data by IoT devices. In addition, the platform collects the data of the physical and mental condition that users determined by Slack.

Figure 2 illustrates an overview of the platform, and Table 1 shows the data to be collected by the platform. The main purpose of this platform is to integrate these data and enable data collection, storage, and access for presenting information for maintaining health. The details of each application are described below.

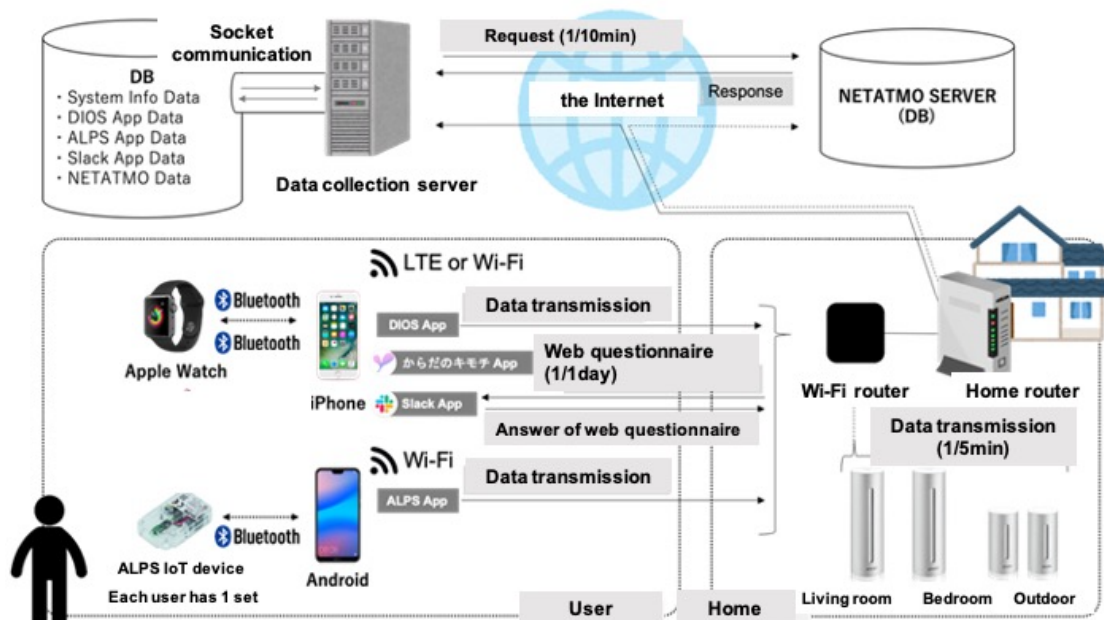


Figure 2 System overview.

Table 1 Collection data list in this study

Data types	Device	Data collection interval
Acceleration	iPhone	1seconds
Location information	iPhone	Irregular
Number of steps	iPhone	Minimum 10 seconds
Battery capacity	iPhone	1 minute
Heart rate	iPhone / Apple Watch	Irregular (minimum seconds)
Magnetic 3 axes	ALPS sensor	1 second
Acceleration 3 axes	ALPS sensor	1 second
UV	ALPS sensor	1 second
Ambient light	ALPS sensor	1 second
Temperature	ALPS sensor	1 second
Humidity	ALPS sensor	1 second
Barometric pressure	ALPS sensor	1 second
Voltage	ALPS sensor	1 second
Living room temperature	NETATMO Master unit	About 5 minutes
Living room humidity	NETATMO Master unit	About 5 minutes
CO2 in the living room	NETATMO Master unit	About 5 minutes
Living room noise	NETATMO Master unit	About 5 minutes
Living room air pressure	NETATMO Master unit	About 5 minutes
Bedroom temperature	NETATMO Master unit	About 5 minutes
Bedroom humidity	NETATMO Master unit	About 5 minutes
CO2 in the bedroom	NETATMO Master unit	About 5 minutes
Bedroom noise	NETATMO Master unit	About 5 minutes
Bedroom air pressure	NETATMO Master unit	About 5 minutes
Outdoor temperature 1 (living)	NETATMO Cordless handset	About 5 minutes
Outdoor humidity 1 (living)	NETATMO Cordless handset	About 5 minutes
Outdoor temperature 2 (bedroom)	NETATMO Cordless handset	About 5 minutes
Outdoor humidity 2 (bedroom)	NETATMO Cordless handset	About 5 minutes
Questionnaire data of Physical and mental condition	Slack	Once a day

The platform collects 29 types of data. For that, the platform has function of a) collecting data from the IoT device, which named NETATMO, with its own database from the Web API, b) collecting data from smartphone, which connects wearable devices, that Apple Watch was selected, and IoT devices, that ALPS IoT device was selected.

The platform consists of the following three applications; data collection server CIOAppS (Central Information Operation Application Sever), Apple Watch and iPhone's data collection server DIOSApp (Data Information Operation Service Application), ALPS IoT Smart Modul3 and Android smartphone's compatible application AIOSApp (ALPS Information Operation Service Application). We adopted the Apple Watch as a device for collecting body data and the ALPS sensor (IoT Smart Module sensor network module [11]) as a device for collecting the surrounding environment of the body. In addition, NETATMO, which is one of IoT devices, was adopted to collect indoor environment data [12]. The details of three applications are described below.

### 3.1 CIOAppS

The function of this application is to store the data sent from DIOS App of iPhone and Apple Watch, the data sent from AIOS App, the answer data of Slack questionnaire, and the NETATMO data in the database. It also has a function to send questionnaires to Slack and a function to collect NETATMO data on a regular basis. In other words, this application plays a role of managing data storage, processing, analysis, and presentation when connecting to other devices for data collection as an IoT platform. The functions of this application are as follows.

- Access restrictions as a login function for data collectors
- Store data sent from DIOS App in DB
- Store data sent from AIOS App in DB
- A function that periodically sends a questionnaire to a specific channel using Slack and periodically sends the response data to the server.
- A function to acquire the data (temperature, humidity, illuminance, CO2, noise) of the master unit and slave unit measured by the environmental data by NETATMO at specified intervals and store them in the database

MongoDB, which is a document-oriented database, was used for this application, and the data is saved in BSON format, which is a binary JSON. One record is divided and saved for each day based on UTC time 0:00:00 and for each data type (DIOS, ALPS, Slack, NETATMO). Table 2 shows the development environment of CIOAppS. Figure 3 shows the administrator screen of this application.

Table 2 Development environment of CIOAppS.

Server side	
OS	Mac OS
development language	Node.js (v10.18.1)
Major libraries	express, ejs, crypto, Socket.IO
Database	MongoDB
client	
OS	Mac OS
development language	HTML, CSS, (SCSS), JavaScript

Major libraries	Vue.js, Bootstrap-Vue, Core UI, chart.js, Axios
-----------------	---

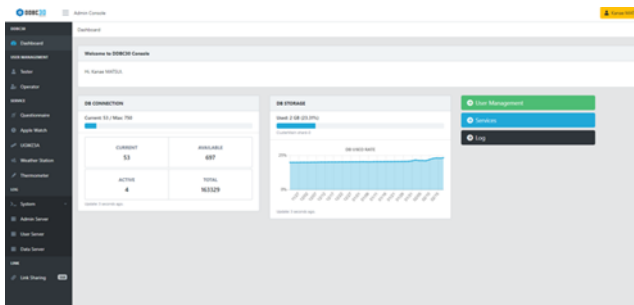


Figure 3 Screenshot of CIO AppS admin page.

Accounts for the experiment participants were created in Slack account created by us in advance. We also downloaded the application to iPhone and set it so that the participants can log in with the created account. For these accounts, the questionnaire is sent from the bot in Slack at 19:00 on time. Figure 4 shows the questionnaire posted to Slack. The questionnaire asked the participants about their daily physical and mental sufficiency, and if there is a reason, they are asked to fill in the free response column.

2020年 2月 18日のアンケート

Q1. 今日の身体の調子はどうでしたか？  
選択してください

Q2. 身体の調子について気になる点があれば記入してください (任意)  
例) 朝から頭痛がした。腰の痛みを感じた。...

Q3. 今日の心の調子はどうでしたか？  
選択してください

Q4. 心の調子について気になる点があれば記入してください (任意)  
例) 一日気分よく過ごせた。あまり気持ちが乗らず。充足感を得られなかった。...

自由記述欄

① 実験実験アンケートについてもっと詳しく キャンセル 回答する

Q1. Please select your physical condition from 7 scales

Q2. Please write the reason of your selection of Q1.

Q3. Please select your mental condition from 7 scales

Q4. Please write the reason of your selection of Q3.

Figure 4 Screenshot of Slack based questionnaire

### 3.2 DIOSApp

This application acquires the data from iPhone and Apple Watch. The application is necessary to install it in advance in the iPhone before distributed to the experiment participants. The application has a function to periodically send latitude, longitude, horizontal accuracy, vertical accuracy, acceleration (X, Y, Z) data measured by iPhone, and BMP data from Apple Watch to a specified server. The data collected by the application is sent to the CIOAppS database. For data transmission, we signed up for SORACOM Air for Cellular and plugged it into the iPhone in advance. Table 3 shows the development environment of DIOS App.

Figure 5 shows a diagram explaining how to use this application. Since DIOSApp starts in the background and acquires data, the experiment participants start the application

by tapping the icon and periodically check the application's movement.

Table 3 Development environment of DIOS App.

OS	Mac OS
development language	Objective-C, Node.js (v10.18.1)
Major libraries	React-Native (v0.59.10) , Native-Base, React-Native-Push-Notification
Integrated development environment	Xcode



Figure 5 Screenshots of CIOAppS.

### 3.3 AIOSApp

This application is compatible with ALPS IoT Smart Module and has a function to periodically send geomagnetic field (X, Y, Z), barometric pressure, temperature, humidity, and UV data measured from compatible devices to a specified server. The application must be installed in advance in the Android device before the experiment.

The data of magnetic 3-axis, acceleration 3-axis, UV, ambient light, temperature, humidity, barometric pressure, battery remaining amount (voltage) data are measured by ALPS IoT Smart Module paired with Android terminal (HUAWEI nova lite2 this time). Also, the application has a function to periodically send to a specified server, CIOAppS database. Table 4 shows the development environment of AIOS App. Figure 6 shows a screen shot of this application installed on an Android device and screen transitions when used.

Table 4 Development environment of AIOSApp.

OS	Mac OS
development language	Android-Java
Integrated development environment	ALPS sensor SDK for evaluation

Using the above three applications, we conducted a demonstration experiment to collect physical and mental state data in households of the child-rearing generation. The details are described below.





Figure 6 Screenshots of AIOsApp.

## 4 EXPERIMENT

The following items were carried out in preparation for the experiment using the above application.

1. Device management
2. Create an account on each device

In item (1), the management number and barcode of each device (NETATMO master unit, slave unit, iPhone, Apple Watch, Android terminal, IoT Smart Module, BUFFALO small router) used. In addition, the account (e-mail address and password) created in item (2) was issued to each device and set to be usable. Tables 5 shows the contents of the kit sent to each household. Each kit was sent to fathers and mothers in families with small child or children.

Table 5 List of kit packing contents

No.	Contents	No.	Contents
1	BUFFALO Small Router	10	Lightning Earphones
2	USB Micro-B cable for power supply	11	AC adapter (5W)
3	Power supply AC adapter	12	NETATMO main unit For living room
4	LAN cable	13	NETATMO main unit adapter
5	Apple Watch body (with band)	14	NETATMO outdoor module
6	Dedicated charging cable	15	Android smartphone body
7	AC adapter (5W)	16	Android Smartphone Charger
8	Apple iPhone	17	ALPS sensor
9	Lightning Charging Cable	18	Demonstration Experiment Kit Bag

As shown in Fig. 7, the devices for which the above preparations completed and made into kits and were distributed to the experiment participants.

The wearable device was requested to be worn from the time of waking up in the morning until going to bed at night, and to carry a smartphone. This is because the wearable device is paired with the smartphone. For Slack use, we requested the response to the questionnaire after 19:00, when there is a regular transmission. These requests were not binding and were limited to voluntary participation.



Figure 7 Layout of the figure object.

Table 6 Participants of the feasible study.

Household ID	Household composition
A	Married couple (male 30s, female 30s), 2 children
B	Married couple (male 30s, female 30s), 2 children
C	Married couple (male 40s, female 40s), 3 children
D	Married couple (male 50s, female 40s), 2 children
E	Married couple (male 50s, female 30s), 2 children

### 4.1 Experimental period

The experiment period was from February 6<sup>th</sup> to March 6<sup>th</sup>, 2020. In order to secure data collection for one month. The data collection period was not uniform because there are household differences in the setup of the delivery kit.

### 4.2 Feedback to the participants

Although the emphasis of this study is on the usefulness of the platform, the main goal of the study subject is “to maintain the health of the child-rearing generation”, and data collection for that purpose. For this reason and to encourage motivation to participate in the experiment, we provided the following feedback through Slack. The feedback sets four types of factors that hinder health maintenance, determine whether or not health maintenance is possible from the data and provides information for individuals based on the data. The factors that hinder health maintenance are “insufficient amount of exercise”, “insufficient rest”, “problems with living environment”, and “concern about mental fatigue”.

Table 7 shows the judgments and the data on which the judgments are based.

Table 7 A method of feedback for maintaining healthcare.

Subjects	Rationale
Lack of exercise	Acceleration, location information (within 10m), steps, heart rate
Lack of rest	Obtained data as a basis: Acceleration, location information (within 10 m), number of steps, heart rate, Slack questionnaire data
There is a problem with the living environment	A sensor device NETATMO is installed indoors (living room, bedroom) to measure indoor temperature, humidity, illuminance, CO <sub>2</sub> , and noise. Evaluate whether the living environment is comfortable

Concerned about mental fatigue	Survey response data obtained from Slack
--------------------------------	--

When the subjects in the table were confirmed from the data, the feedback would be provided each participant with advice to improve or maintain each situation via Slack.

## 5 EVALUATION

The main purpose of this paper is to collect data showing the physical and mental states of the child-rearing generation, and this chapter evaluates the data collection status.

### 5.1 Physical data

The data described in this section are collected by Apple Watch, ALPS sensor, and NETATMO. Each feature is described below.

#### Apple Watch data

In this device, a stable data acquisition situation was observed from the data. Probably, the watch type and the mounting method are widespread among the participants. The stabilization of data collection in this mounting method will be described in detail in the ALPS sensor data described later, but it was found that it has a great deal to do with data acquisition. However, since the resolution was set to 1 second in order to determine “insufficient momentum” in Table 7, the battery is exhausted, and frequent charging is required. It was found that the collection rate drops significantly in the case of collaborators who do not voluntarily habituate this behavior.

#### ALPS sensor data

In this device, unlike a watch-type wearable device such as the Apple Watch, it is a device that is worn like a necklace from the neck, and it is not a common way to wear it. Figure 8 shows the data obtained from one collaborator.

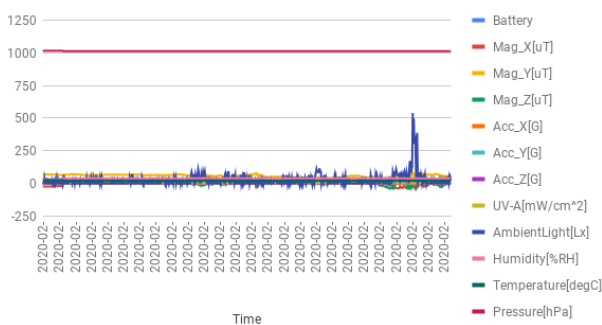


Figure 8 Collected data by ALPS IoT sensor.

Normally, it is expected that each acquired data will draw a larger waveform during daytime work, but the vertical movement of the waveform is small, and the sensor was put in or removed from the chest pocket. It is expected that it was placed.

#### NETATMO data

In this device, stable data acquisition was observed because it was not a wearable type. Since the installation type device only requires the intervention of the participants at the time of setting, it was found that if the initial installation method is correct, data can be collected stably. However, the data collected from household E in Fig. 9 as an example, it was found that the room temperature did not rise to nearly 20°C unless it was around 11:00 am, even in the living room where the family gathered.

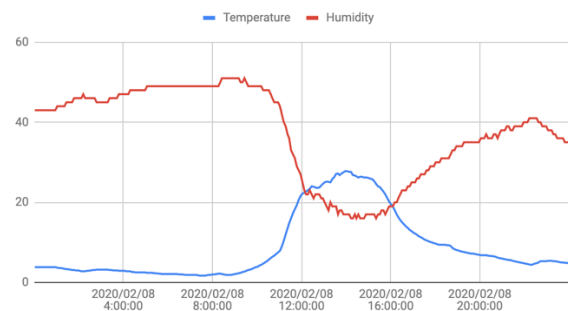


Figure 9 Collected data by NETATMO.

Since the installation location of the environmental sensor installed indoors has a large effect, whether the location where this sensor is placed reflects the typical indoor temperature of the living room, or in a place where the temperature is difficult to reflect, such as inside a shelf. It is necessary to determine whether it was placed and have it set the installation position in a place where the temperature is reflected.

### 5.2 Mental state data

In this data, the recovery rate of the response data to Slack, which is regularly sent from CIOAppS, and the collected response data are evaluated. The recovery rate of this data was about 80%. There were two participants who never responded, but since they were participants from other households, it is not that they did not know how to use Slack, and it is probable that they did not have the motivation to respond.

Among the other participants, the time zone of response and the quality of response were remarkable. The response rate was high among the collaborators whose response time zone had become established, and there was a description in the free description column as to why they chose the scale. Respondents who showed this tendency responded by giving the feedback described in Table 7 and tended to report their own changes in consciousness and behaviors. On the other hand, there was no response even when feedback was sent to the collaborators who responded in a sparse time zone and took actions such as sending answers for several days in their spare time.

### 5.3 Usability of this kit

As described in Chapter 4, since various devices were installed and installed in the demonstration experiment in this study, a post-questionnaire was conducted to see if the results would be a burden to the participants in the experiment. This



item is provided because if the feeling of use is poor, it will not lead to the use of results and will not lead to correct data collection.

In addition, in the free response, one of the participants said that it was difficult to carry around because the total weight of the two smartphones used in this experiment and his own smartphone was about 1 kg. So, it is necessary to reduce the amount of equipment.

## 6 CONSIDERATION

This chapter describes the consideration based on the evaluation results of the previous chapter. Since this paper focused on the construction of the platform and the verification of its usefulness, determined from the situation of the data obtained from the experiment. The function of the platform operated normally, and it was possible to collect and accumulate the data. However, more detailed analysis should be required for observing data quality. The usage status of the devices has a significant effect on the data, so improvement can be requested from the experimental operator such as charging and mounting method.

In addition, considering the response from the feedback conducted this time, if feedback that is determined to be correct for maintaining or improving the physical and mental good condition, becoming positive effects were observed with the platform. We obtained the knowledge that the technical and operational improvement methods of the data collection rate from this experiment and can be reflected in the platform.

## 7 CONCLUSION

The child-rearing generation is expected to play a leading role in economic activities, but due to changes in various life stages such as marriage, child-rearing, and promotion at work, physical and mental illnesses such as lifestyle-related diseases and mental depression due to environmental changes are mental and physical. It is important to be in good health before suffering a major illness from the viewpoint of non-illness, and since the physical and mental condition changes daily. Therefore, we proposed a platform for conducting monitoring. In addition, this monitoring is data collection for presenting information for the purpose of maintaining health based on scientific evidence, and from (1) behavioral changes to maintain data accuracy by monitoring, and (2) acquired data. We aim to realize information presentation that is expected to contribute to the maintenance of health for each individual.

In this paper, the experiment using the platform to verify its availability was carried out for about one month with the participation of five households of the child-rearing generation. Based on the feedback obtained this time in the system construction, we will further refine the information providing according to the situation of the target person from the system refinement and data analysis. In addition, in order to encourage the use of the application developed in this study in other studies and services that are expected to utilize the data collected in this experiment, we will proceed with the development as open source.

## Acknowledgments

This study activity was carried out by the Society 5.0 Realization Study Center Support Project (S004541) by the Ministry of Education, Culture, Sports, Science and Technology.

## References

- [1] Ministry of Health, "Labor and Welfare Issues by age group and generation," URL: <https://www.mhlw.go.jp/content/12404000/000528279.pdf>, only in Japanese (2021-06-12).
- [2] Otsuka R, Tamakoshi K, Yatsuya H, et al. Eating fast leads to obesity: findings based on self-administered questionnaires among middle-aged Japanese men and women. *J Epidemiol* 2006; 16(3): 117–124.
- [3] Ministry of Economy, "Trade and Industry Promotion of health management," [http://www.marutaka-g.co.jp/kenko\\_keiei/pdf/180710kenkoukeiei-gaiyou.pdf](http://www.marutaka-g.co.jp/kenko_keiei/pdf/180710kenkoukeiei-gaiyou.pdf), only in Japanese (2021-06-12).
- [4] REN, Ju, et al. Serving at the edge: A scalable IoT architecture based on transparent computing. *IEEE Network*, 2017, vol. 31, no. 5, p. 96-105.
- [5] JARWAR, Muhammad Aslam; ALI, Sajjad; CHONG, Ilyoung. Exploring web objects enabled data-driven microservices for E-health service provision in IoT environment. In: 2018 International Conference on Information and Communication Technology Convergence (ICTC). IEEE, 2018. p. 112-117.
- [6] MATSUI, Kanae. An information provision system according to residents' indoor comfort preferences for energy conservation.
- [7] FUKUTA, Mio, et al. Proposal for home energy management system to survey individual thermal comfort range for HVAC control with little contribution from users. In: 2015 IEEE 13th International Conference on Industrial Informatics (INDIN). IEEE, 2015. p. 658-663.
- [8] OKOSHI, Tadashi, et al. WellComp 2020: third international workshop on computing for well-being. In: Adjunct Proceedings of the 2020 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2020 ACM International Symposium on Wearable Computers. 2020. p. 671-674.
- [9] VISURI, Aku, et al. Understanding smartphone notifications' user interactions and content importance. *International Journal of Human-Computer Studies*, 2019, vol. 128, p. 72-85.
- [10] "slack" . <https://slack.com/intl/ja-jp/>, (2021-06-12).
- [11] "ALPS Sensor Network Module" . <https://tech.alpsalpine.com/j/products/iotsmart-network/>, (2021-06-12).
- [12] "Measure your environment," URL: <https://www.netatmo.com/en-us/weather>, (2021-06-12).



# Development of a Safety Training System for a Portable Grinder that Combines Virtual Reality with an Actual Tool

Tomoo Inoue\*, Asuki Nakanishi\*\*

\*Faculty of Library, Information and Media Science, University of Tsukuba, Japan

\*\*Graduate School of Library, Information and Media Studies, University of Tsukuba, Japan  
inoue@slis.tsukuba.ac.jp

**Abstract** - Appropriate safety training to workers to prevent labor accidents is crucial in industry. In this research, a safety training system using virtual reality and an actual portable grinder to learn the risks of accidents is proposed. The system is composed of an actual portable grinder, a mechanical device to provide force feedback, and an HMD for visual VR experience. It also has a grinder-type controller to move a virtual grinder for virtual sparking experience. Users can learn the risks by experiencing these virtual accidents visually, auditorily, and haptically.

**Keywords:** virtual reality, safety training, physical computing, grinder.

## 1 INTRODUCTION



Figure 1: A portable grinder (Model: RYOBI G-1061P).

Applying appropriate safety training to workers to prevent labor accidents at industrial sites is widely recognized as being a responsibility for the workers of the business operators.

Meanwhile, in recent years, inexpensive, high resolution, wide viewing angle head mounted display (HMD) has appeared, and safety training method using virtual reality (VR) technology is appearing. In this research, a safety training system that combines VR technology with an actual grinding machine is proposed for preventing labor accident in grinding work by a portable grinder.

A portable grinder is a tool typically used for scraping and polishing metal and has a grinding wheel (Figure 1). A grinder rotates the grinding wheel by which metal is contacted.

The possible accidents regarding a portable grinder and the guideline to prevent them is publicized by Japan

International Training Corporation Organization (JITCO) [1].

## 2 RELATED WORK

### 2.1 Safety Training Systems using VR Technology

A few existing examples of safety training systems to prevent accidents and dangers using VR technology are a safety training system for miners with joysticks and touch panels by Squelch [2], a safety training system for learning safe handling methods by Nakayama et al. [3], and a safety training system to learn dangers when crossing a road for children by McComas et al. [4]. They only provide visual information and accept input devices. In a system by Diez et al. [5] that educates how to deal with a fire, Oculus Rift is used to present a virtual space, and Leap Motion detects the motion of fingers, and a virtual fire extinguisher needs to be operated by the motion of fingers. Although it accepts user's behavior as input, it provides visual information only.

Users cannot feel force, weight and other senses that would increase reality from these VR systems.

### 2.2 Safety Training Systems using an Actual Tool

As a safety training system using only the actual grinder without using VR technology such as a display that presents virtual space to the user, Nihon Mechanic Co. Ltd developed a safety training system to prevent the accident of the actual portable grinder.

It consists of a stationary machine, a grinder fixed to the fixture of the machine, a wooden stick as an object to be grinded, and a balloon. In the system, the reaction force can be given to the grinder by air pressure. The grinder is thus repelled strongly to break the balloon placed in the opposite end of the equipment. The sound of the balloon breaking is made to surprise the user.

Although there are safety training systems that enable simulated accident experiences using actual equipment, such as this system, the appearance differs greatly from actual workplaces.

### 2.3 Safety Training System Combining VR Technology and an Actual Tool

There have been simulators that provide other sensations than visual stimulation. However, most simulators providing physical sensations are not portable. A small-sized excavator simulator for safety training still provide an operator's seat with motion stimulation, and is not portable, though it was reduced in size significantly from the previous one [6].

A safety training system that provides physical stimulation but is portable is not easy to implement in many cases. If it is portable, however, it is desirable because it eliminates restriction on the training locations. In this study, we propose a portable safety training system for a Portable grinder that uses an HMD, a laptop computer, and a few more devices.

### 3 PROPOSED SYSTEM

#### 3.1 System Overview

In this study, we propose a safety training system that combines VR technology and an actual machine (AM) for workers who use portable grinders, to support the learning of knowledge to prevent accidents.

A complete set of the proposed system consists of an HMD, a location sensor using an infrared camera, an actual grinder with the sensor attached (Figure 2), a grinder-type controller (Figure 3), and a pseudo-repulsion-experience machine (Figure 4). In the proposed system, the position and orientation of the grinder is acquired by the attached position sensor, and a virtual grinder placed in the simulated work space is synchronized with it (Figure 5). By combining VR technology with AMs, users can experience a higher sense of reality than with systems using only VR technology or only AMs. The system aims to make users aware of the dangers of handling a grinder in an inappropriate manner through virtual accident experience under such condition.

The proposed system offers two training scenarios, "repulsion experience" and "spark experience," which are described in the following sub-sections.

#### 3.2 Repulsion Experience

In the grinding operation using a portable grinder, it is known that when the grinding wheel of a grinder collides with the object to be ground, the grinder can be bounced by the reaction, causing the risk of injuring the body of the user or the surrounding persons. The "repulsion experience" scenario is aimed to make users experience virtual repulsion of a grinder and let them learn the dangers thus caused.

In the repulsion experience scenario, a user starts from wearing the HMD and holding the grinder attached to the pseudo-repulsion-experience machine. When the user moves the actual grinder to the right, the virtual grinder that is synchronized with the actual grinder also moves to the right. It can touch the cylindrical object, which is to be ground. When the grinding wheel of the virtual grinder touches the object, particles like a spark is presented (Figure 6). When it touches the object in a bad manner, the pseudo-repulsion-experience machine reacts to it, making the grinder physically bounced off to the left.



Figure 2: For the repulsion experience, Oculus touch is attached to the upper right part of the actual grinder.



Figure 3: Grinder-type controller.



Figure 4: Pseudo-repulsion-experience machine.





Figure 5: Virtual grinder.



Figure 7: Target object.

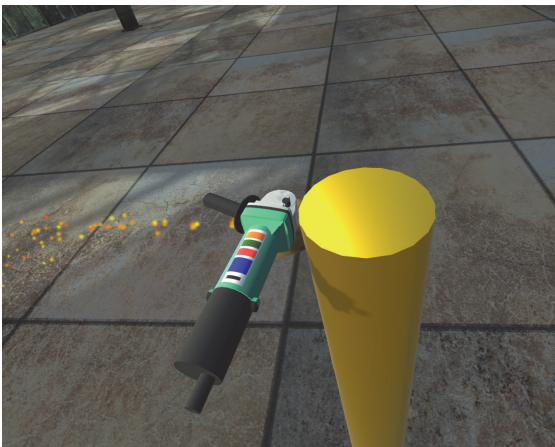


Figure 6: Cylindrical object is contacted with a grinding wheel.



Figure 8: Spark particles.

Unity in diameter and 0.8 Unity in height, where “Unity” refers to a unit in the virtual space that a person perceives as 1 meter in the physical world. The virtual grinder is produced with reference to an actual grinder (Figure 5). When the grinding wheel of the virtual grinder touches the cylindrical object, sparking particles are presented (Figure 6). When it touches the object and still moves to the right (to 16% (heuristically determined by an expert) of the grinding wheel in diameter overlaps with the object), the pseudo-repulsion-experience machine is activated.

### 3.3 Hardware of Repulsion Experience

The repulsion experience scenario uses the HMD (Oculus Rift), a laptop computer, and the pseudo-repulsion-experience machine with the portable grinder attached (Figure 4).

This machine has an iron spring for presenting bouncing force, a metal fitting for holding the spring when extended, Arduino to controls it, and a fixing tool of the grinder with its base on a rail and attached to an end of the spring. When electric current flows to the circuit via Arduino, the metal fitting is released and the spring returns to its original length, causing the repulsion of the grinder to the left end of the rail. To obtain the position and angle of the grinder, Oculus Touch is attached to the grinder (Figure 2).

### 3.4 Software of Repulsion Experience

The virtual space for the repulsion experience scenario was made by Unity 5.6.3f1 Personal (64bit). The virtual space used as a basis is "Steel Mill Warehouse" in the Unity Asset Store [7]. The size of the cylindrical object is 0.2

### 3.5 Spark Experience

When a metal is polished by a portable grinder, chips and dust fly as a spark. The spark experience scenario is for a user to experience and learn how a spark flies, how a grinder can be operated safely to control a spark direction.

The target object to polish is placed in the virtual space (Figure 7). The user operates the virtual grinder by moving a grinder-type controller freely. When the grinding wheel of the virtual grinder meets the target object to polish, spark particles appear (Figure 8).

### 3.6 Hardware of Spark Experience

The spark experience scenario uses an Oculus Rift HMD, a laptop computer, and a grinder-type controller. The grinder-type controller is a simple mock-up with a body and a handle similar to an actual grinder, with Oculus Touch attached for position sensing (Figure 3).

### 3.7 Software of Spark Experience

The virtual space for the repulsion experience scenario is also used in the spark experience scenario. The target object to polish is a metal shelf with 1 Unity in width and 2.3 Unity in height (Figure 7).

<https://assetstore.unity.com/packages/3d/environments/industrial/steel-mill-warehouse-17718>.

## 4 CONCLUDING REMARKS

A safety training system for a portable grinder that combines VR with an actual tool was presented. While this system maintains the merit of combining VR and physical object to give user safe yet realistic experiences, it is designed to be portable enough to be used in a training course held in various places other than a single place where the large system is installed. The system offers two typical scenarios for experience accidents; repulsion of a grinder and sparking. They are realized by implementing specific devices such as a pseudo-repulsion-experience machine and a grinder-type controller, but other than these they can use common virtual space and equipment.

The evaluation of the system will be our future work.

## ACKNOWLEDGEMENTS

This research was partially supported by the collaborative research project “Research on VR equipment for the training of occupational accident prevention” with Nihon Mechanic Co. Ltd.

## REFERENCES

- [1] Japan International Training Corporation Organization. Retrieved January 15, 2017 from [https://www.jitco.or.jp/download/data/kensakuban\\_sai\\_gaibousi.pdf](https://www.jitco.or.jp/download/data/kensakuban_sai_gaibousi.pdf)
- [2] Andrew Squelch. 2001. Virtual Reality for mine safety training system in south Africa. The journal of the South Africa Institute of Mining and Metallurgy. 209-216.
- [3] Shoji Nakayama, Ge Jin. 2015. Safety Training Enhancing Outcomes Through Virtual Environment. American Society of Safety Engineers. 60,2. 34-38.
- [4] Joan McComas, Morag Mackay, Jayne Pivik. 2004. Effectiveness of Virtual Reality for Teaching Pedestrian Safety. CyberPsychology & Behavior. 5,3. 185-190.
- [5] Helen V.Diez, Sara Garcia, Andoni Mujikam Altor Moreno, David Oyarzun. 2016. Virtual Training of fire warden through immersive 3D environments. Proceeding Web3D '16 Proceedings of the 21st International Conference on Web3D Technology. 43-50.
- [6] Kiyoshi Fukaya, Takahiro Nakamura. 2005. Development of a Small Sized Simulator to Simulated Hazards in Excavator Accident. Specific Research Report of the National Institute of Industrial Safety. 31-40.
- [7] Aron Versteeg. Steel Mill Warehouse. Retrieved 30 October, 2021 from

# Development of a benchmark system on power-related control by BACnet/IP

Kohei Miyazawa\*, Tetsuya Yokotani\*\*, and Hiroaki Mukai\*\*\*

\* Electrical and Electronic Science Engineering, Kanazawa Institute of Technology, Japan  
c6100786@planet.kanazawa-it.ac.jp

\*\* Engineering department, Kanazawa Institute of Technology, Japan  
yokotani@neptune.kanazawa-it.ac.jp

\*\*\* Engineering department, Kanazawa Institute of Technology, Japan  
mukai.hiroaki@neptune.kanazawa-it.ac.jp

**Abstract** – BACnet/IP has been applied to the interconnection between end devices and monitoring centers for the management and control of buildings. However, this system's security risks are a concern. In building monitoring and control, each device is controlled remotely, and unauthorized access causes major accidents. This study describes an evaluation platform to address these issues. This evaluation platform emulates the case of unauthorized operation of power switchgear via BACnet/IP and analyzes the impact on each device connected to the evaluation platform. Additionally, the platform implements a mechanism to detect and shut down unauthorized operations. The evaluation platform is connected to a single-phase 100 V and three-phase 200 V power supply, and the built-in programming logic controller turns it on and off to monitor the effects of the connected devices (temperature rise, arcing). Furthermore, an algorithm for detecting unauthorized access can be implemented. The algorithm makes it possible to efficiently analyze the effects of various unauthorized access and study countermeasures without using actual equipment.

**Keywords:** Building automation system, BACnet, Cyber physical system, Cyber physical security

## 1 Introduction

As Society 5.0, which is proposed to achieve both economic development and resolution of social issues through the advanced integration of cyber and physical space, the internet of things (IoT), which connects things to the Internet, is becoming more widespread and contributing to users by creating various values [1]. The building automation system (BAS) is a common application of IoT, which integrates the management of lighting, air conditioning, security, elevators, and other equipment across a network to achieve monitoring and control within a building [2]. This system employs the building automation and control networking protocol (BACnet/IP), an open protocol that allows for the control of incompatible equipment from different vendors, contributing to improved convenience and safety for users. However, from a security viewpoint, the interconnectivity of the network makes it easy to carry out attacks between networks. The occurrence of unauthorized behavior that affects the physical space because of an attack may cause irreparable damage to the building equipment [3]. Nevertheless, there is currently no physical platform for evaluating unauthorized behavior in building equipment, and no method for building an evaluation platform has been

established.

This study aims to build a platform to evaluate the impact of unauthorized access and malfunction on building equipment in the BACnet environment as a part of the cross-ministerial strategic innovation promotion program 2 (SIP2) project led by the Cabinet Office [4]. Specifically, we assume that the typical control of the switchgear ON/OFF is a malfunction caused by an attack and attacks the equipment in a simulated physical environment. We then built a platform to evaluate the effects of heat rise and electrical wear that may occur during the operation of the equipment.

## 2 Motivation

In this study, we discuss the reason for switching ON/OFF as an operation that impacts the equipment in the building. The operation of each equipment in a building is performed by opening and closing the load current using a switchgear in the distribution board [5]. This means that the equipment connected to the BACnet/IP network can control the switchgear from the central monitoring device. In other words, existing building control systems with interconnectivity can be controlled by malicious users through remote control interfaces, and physical actions can be taken.

How do physical actions affect the equipment? An air conditioning system is a typical example of building equipment. Generally, it is in constant operation when people are in the building, and it is not expected to be turned on and off frequently in a short time. If the power is turned on and off frequently for a short time, the fans and compressors in the air-conditioning system and the capacitors in the inverter will be damaged by the inrush current, exceeding the specified capacity, or the temperature will rise abnormally because of Joule heating [6]. Additionally, the electromagnetic switch on the power supply side may also catch fire because of a short circuit caused by wear and tear of the switching block, or because of welding of the blockage caused by arc discharge between the contacts. This highlights the fact that switching is a cause of various fires, and it is necessary to recognize that switching can be a serious threat.

## 3 Overview of BACnet/IP

This section summarizes BACnet/IP, which has been applied to building management as a typical worldwide protocol.



### 3.1 BACnet/IP feature

BACnet is a communication standard for building networks that became an ISO standard in 1995, as ANSI ASHRAE Standard 135-1995. Even if building equipment such as lighting, air-conditioning, security, and elevators have their own specifications, they can be interconnected through a common interface called the BACnet protocol, which enables the monitoring and control of large-scale buildings [7,8]. BACnet/IP is a technology that realizes the higher-level protocols of the BACnet standard on an IP network (Fig.1). A BACnet network is composed of a collection of devices that use the IP protocol for communication, which means that various BACnet-enabled devices are connected to the ethernet commonly used in all equipment [9]. Using the existing network, the interoperability between systems and the cost of configuring the network infrastructure can be reduced [10]. Devices in BACnet use standard networks and interface cards; Category 5 cables, which BACnet/IP uses for communication, are available in almost all areas of the building. All these lead to cost savings in terms of initial installation, as well as ongoing maintenance and management of the BAS network.

However, it is necessary to understand the issues associated with using BACnet/IP. One of the most critical issues is the security aspect; because BACnet/IP is a standardized and interoperable network, it is not suitable for use in unprotected networks. Therefore, the defense of building management system networks by BACnet/IP on IP networks is essential.

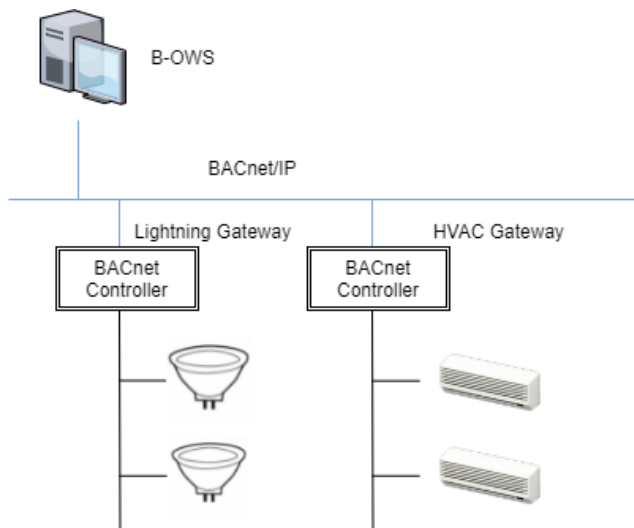


Fig.1 An example of configuration in Building management system

### 3.2 BACnet object

The target data in BACnet are represented by an abstract concept called an object, and the equipment connected to the network is modeled as a collection of objects. Multiple attribute values called properties define BACnet objects. The data possessed by these properties are sent to BACnet devices for processing when they are generated or requested [11].

BACnet objects are classified into basic input/output, device properties, notification functions, life safety, composite functions, file information exchange, and others. The ASHRAE 135-2012 BACnet standard, which is the focus of this study, specifies 54 types of BACnet objects, the most important of which is the binary output object [12]. The binary output object is one of the most used objects to control the starting and stopping of equipment, and it is controlled by writing a binary value to the present value property that manages the current value belonging to the binary output object.

## 4 Proposal on evaluation platform

Each piece of equipment in a typical building distributes power through distribution boards, which are equipped with switches to prevent accidents such as fires. When BACnet is used for building control, the distribution boards can be controlled over a network. In many existing building control systems, it is possible for a malicious user to illegally control the system through a remote-control interface and cause physical abnormalities. Therefore, to verify the validity of the unauthorized behavior patterns, we intentionally cause physical effects on the equipment.

In this evaluation platform, a BACnet unit with a built-in programmable logic controller (PLC) sends high-frequency ON/OFF signals to the switchgear. It then detects the heat rise of the equipment that is turned on and off and determines whether the operation is unauthorized. Additionally, the possibility of a standard operation is examined owing to contact wear of the switch body caused by high-frequency ON/OFF and welding by arc discharge.

### 4.1 Network configuration

The network configuration for applying BACnet is shown in Fig.2.

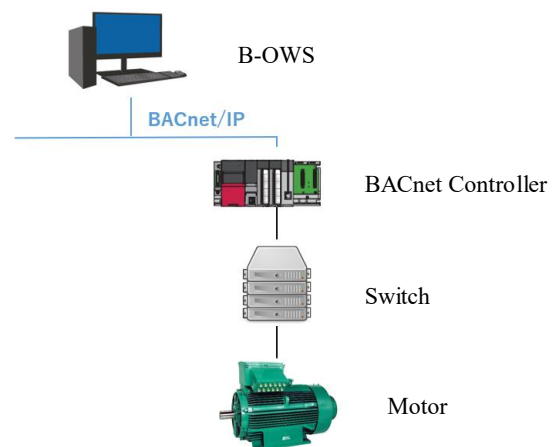


Fig.2 Network Configuration

As shown in Fig.2, B-OWS, the central monitoring device, and the BACnet controller are connected by BACnet/IP, and the BACnet controller and equipment are electrically connected. When the B-OWS sends ON/OFF commands to the BACnet controller, current flows from the switchgear to the equipment, and the equipment o

perates. In this study, we use an inverter motor as an example of a terminating resistor to construct the environment and evaluate the attack pattern.

## 4.2 Hardware configuration

The evaluation platform consists of a BACnet controller, power supply, switchgear, circuit breaker, monitoring panel, and other peripherals. The sequence of the hardware configuration is shown in Fig.3, and its appearance is shown in Fig.4.

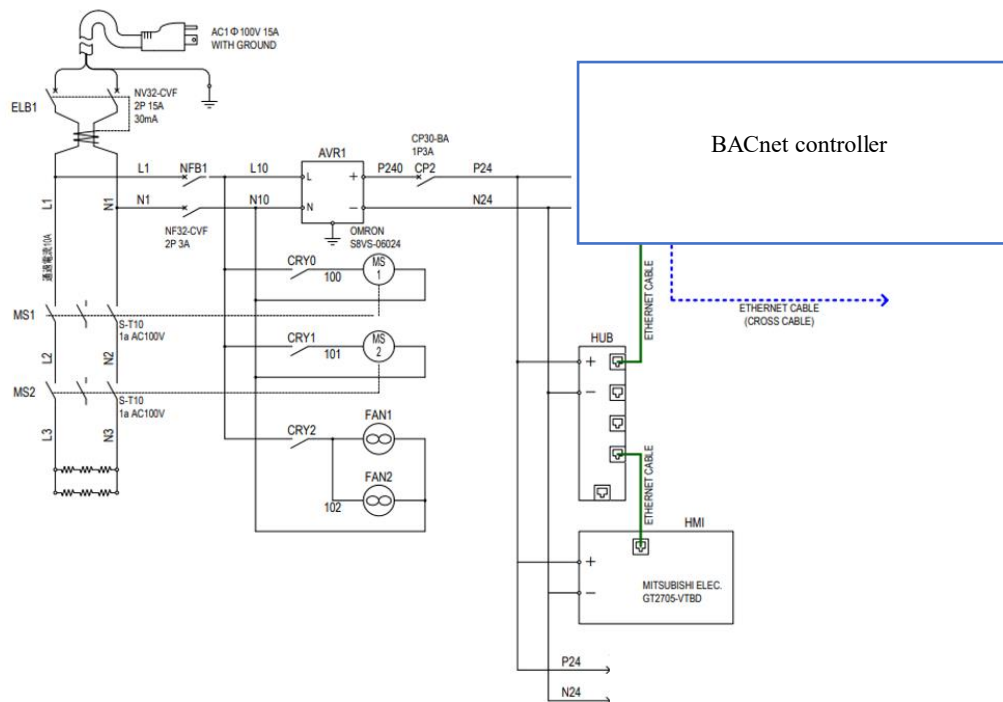


Fig.3 Sequence Diagram of Platform



Fig.4 Appearance of the evaluation platform

In this evaluation platform, two evaluation environments, one with single-phase 100 V and the other with three-phase 200 V, were constructed to evaluate various types of equipment. The 5VDC rectified from these power supplies is input to the BACnet unit to operate the s

witchgear. Additionally, breakers were installed on the primary and secondary sides as fire prevention measures against electrical leakage because a large current is expected to flow when a large-capacity load is installed at the end of the electromagnetic contact.

The BACnet controller with a built-in PLC for BACnet communication is shown in Fig.5. The PLC is connected to the B-OWS using ethernet shown in U1 of Fig.5, and the data between devices are sent and received by registering objects in the BACnet unit of U2. When the property is changed, the transistor output unit shown in U4 turns on and off the switchgear, which is the load. Furthermore, the inter-channel insulation thermocouple input unit shown in U5 inputs the heat data of the equipment to the PLC to evaluate the effect of temperature rise.

PLC control makes it possible to change the increase or decrease of the equipment and the control logic according to the location, creating a highly flexible environment that can be adapted to the situation.

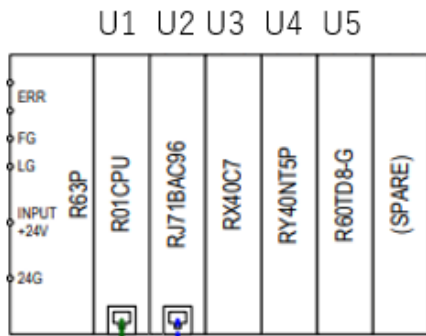


Fig.5 PLC constitution

### 4.3 Safety circuit

Because the test is expected to handle voltage, four safety circuits are used to prepare for the worst-case scenario. The first is overcurrent protection and earth leakage protection with a primary breaker. When the switchgear is unplugged, the breaker is activated to forcibly stop the flow of current. The second button is an emergency stop button, which shuts down the power supply without unplugging the outlet or tripping the internal breaker. Finally, a bimetal switch is used to prevent double-crossing. When a certain amount of heat is generated, the bimetal deforms and physically releases the circuit. Additionally, the PLC monitors the temperature and performs safety control (this will be shown later).

## 5 Control by PLC

The PLC alone installed in the BACnet controller cannot send signals to the switchgear or read the analog values of the temperature sensor to evaluate the temperature rise. Therefore, it is necessary to develop a control program for PLC.

### 5.1 Ladder programming in PLC

The instructions to the PLC are performed using ladder programming, which symbolizes a relay circuit. To write instructions, it is first necessary to register a binary output object from the configuration function installed in the BACnet unit. Fig.6 shows the configuration of the screen. In this evaluation platform, the contents of the binary output object are allocated to the 256 free areas of the buffer memory. Furthermore, a program is created to access the properties allocated to the buffer memory based on the format of the buffer memory of the binary output object.

Specify the area of the buffer memory to which the binary output object is allocated. The value allocated to the present value property is changed by adding the binary value at the end of the bit string, as shown in Table.1, which is the format for specifying the Mitsubishi BACnet unit used in this study. The operation is as shown in Fig.7 and each time the value of property written from B-OWS to the buffer memory in the BACnet unit is changed, the data are read into the PLC main body and the current output signal is sent to the electromagnetic switch according to the value.

InstanceNo	Qty	BufferMemoryAddress
11	1	256 x
<input type="button" value="Add"/> <input type="button" value="Cancel"/>		

Fig.6 Engineering Tool

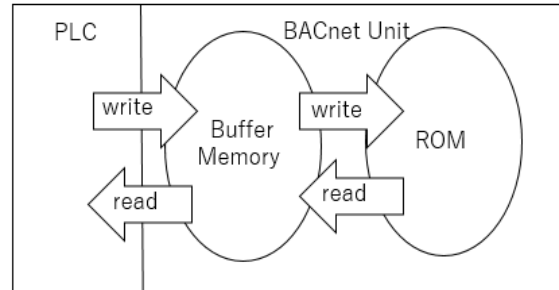


Fig.7 BACnet movement

Table.1 Offset of Binary Output

+ 2	PresentValue	b15-b1		unused
		b0	PresentValue	0:Inactive, 1:Active

### 5.2 Control of evaluation platform

In this study, we assume an attack in which the switchgear is turned on and off at high frequency, so we use a timer instruction to switch the value stored in the present value property between 0 and 1 in a cycle of 100 ms at the shortest, and use a loop instruction to repeat the ON/OFF operation.

Fig.8 shows an example of the ladder program that we created. When the time elapses, the timer contact is turned on, and the value is written into buffer memory 258, which is the offset value of the present value property. When the time elapsed, the timer in D36 was turned off, and the next value was assigned to D38 and the above process loops.



Fig.8 Ladder diagram

### 5.3 Operation of the evaluation platform

The operation panel of the evaluation platform is shown in Fig.9, and an example of the ladder for operation is shown in Fig.10. The internal relay M100 is activated by touching 10212 (internal relay M101). When M100 is turned on, the value stored in the present value property is written to the buffer memory in the BACnet unit, and the switch is activated (Fig.8). Originally, it was necessary to rewrite the value directly into the PLC to change the property value of an object. However, by installing an operation panel, a GUI-based visual operation is possible.

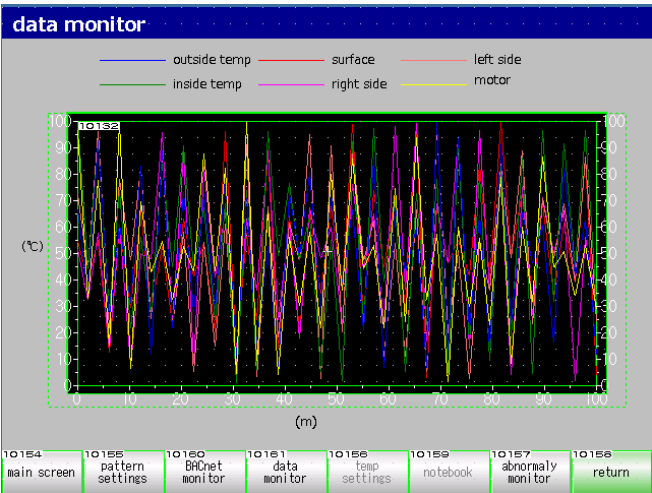


Fig.9 Operation screen

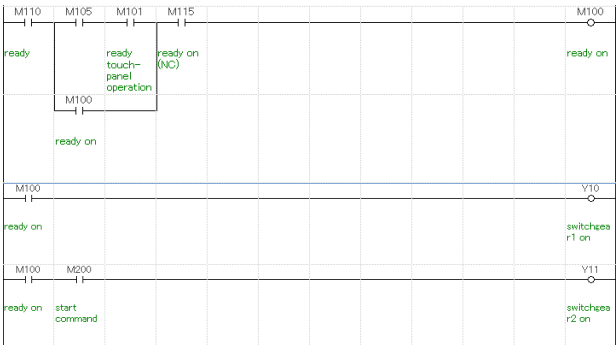


Fig.10 Operation program

5.4 Safety control

We built a system that detects abnormal temperature rise in switchgear and equipment and sends a stop signal from the PLC side. Fig.11 depicts the operation panel for setting the upper temperature limit, while Fig.12 depicts an example of the control system ladder. The inter-channel insulation thermocouple unit shown in u3 of Fig.5 inputs the temperature data to the data registry D0 defined in Fig.11. The upper temperature limit set in Fig.9 is then compared with the contact-type comparison command to release the circuit between the PLC and the switchgear.

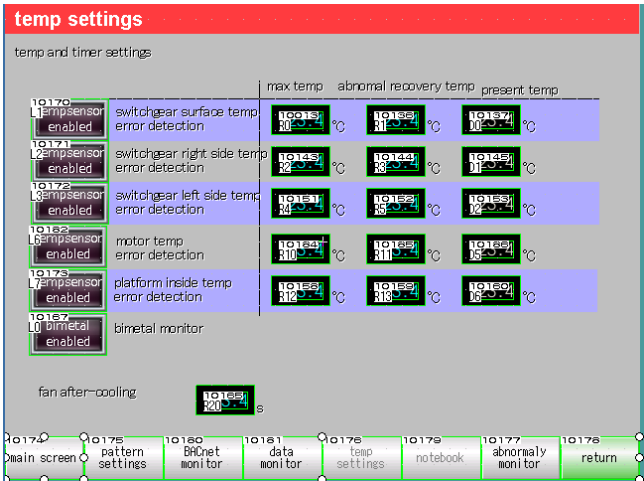


Fig.11 Temperature setting screen

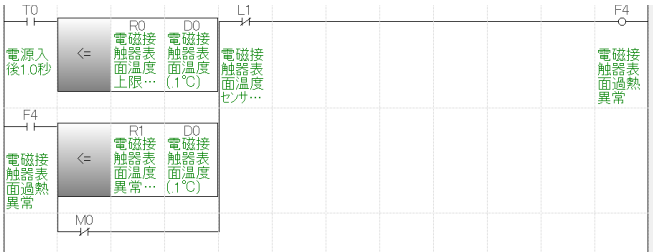


Fig.12 Temperature evaluation program

6 Evaluation method

The evaluation platform targets the temperature rise of the inductor motor, which is the termination resistance of the switchgear, and the effect on the equipment when the switchgear reaches the end of its life.

Fig.13 shows the appearance of the evaluation platform when the temperature data from the built-in temperature sensor are being measured.

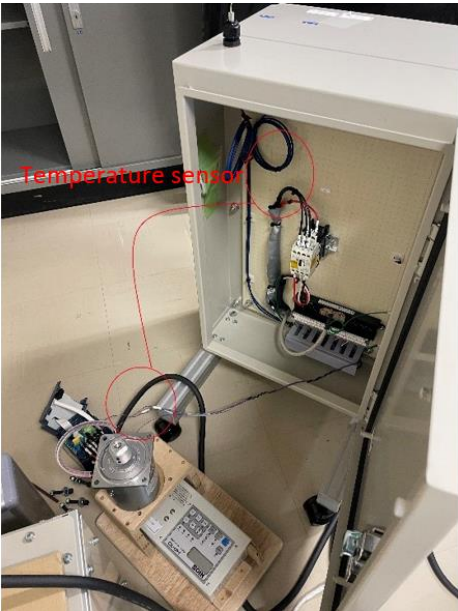


Fig.13 Temperature sensor

The graph shown in Fig.14 displays the temperature data, with time on the X-axis and temperature on the Y-axis, showing temperature changes at one-minute intervals. This evaluation platform has six temperature measurement sensors, each of which has a different color.

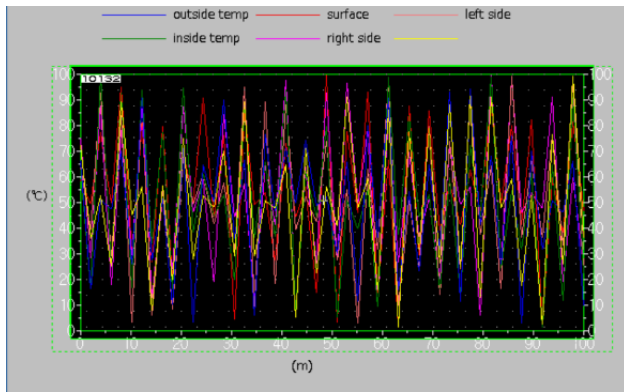


Fig.14 An example of measured data

As shown in Fig.15, the temperature data at 5-second intervals are sampled in csv data on an SD card triggered by operating the evaluation platform, and the data can be extracted as needed.

	A	B	C	D	E	F	G
1	:GT2K_LO	0					
2	:LOGGING	1					
3	:LOGGING TEMP LOG						
4	:SERIAL_I	12					
5	:DEVICE_I	5					
6	:RECORD_I	107					
7	:DATE_OF YYYY/MM/DD hh:mm:ss						
8	:LOCAL_T GMT 00:00						
9	:TIME_INF_ORDER						
10	:DEV_CON SURFACE RIGHT SIDE LEFT SIDE INSIDE TEM MOTOR TEMP(0.1°C)						
11	:DEV_TYP BIN16	BIN16	BIN16	BIN16	BIN16		
12	:DISP_TYP DEC	DEC	DEC	DEC	DEC		
13	:DEV_SIZE	1	1	1	1	1	
14	#####	216	218	218	214	219	
15	#####	216	218	218	215	219	
16	#####	217	218	218	215	220	
17	#####	216	218	218	215	220	
18	#####	216	219	218	215	220	
19	#####	216	219	218	215	220	
20	#####	216	219	218	215	220	
21	#####	215	219	218	215	220	
22	#####	215	219	218	215	220	
23	#####	215	219	218	215	219	
24	#####	215	219	218	215	220	

Fig.15 Measured data of excel

## 7 Conclusions

In this study, we developed a platform for evaluating the impact of unauthorized access to building equipment in the BACnet network when the switchgear is turned on and off frequently. This allows us to efficiently analyze the impact of various unauthorized accesses and to study countermeasures without using actual equipment.

## References

- [1] “Society 5.0”, [www8.cao.go.jp/cstp/english/society5\\_0/index.html](http://www8.cao.go.jp/cstp/english/society5_0/index.html) (2020)
- [2] T. Masuko, “Trends in BACnet to the IoT Era”, The Journal of the Institute of Electrical Installation Engineers of Japan, Vol.37, No.3, pp.156-159 (2017)

[3] DG. Holmberg, “BACnet Wide Area Network Security Threat Assessment”, Nist Interagency/Internal Report (NISTIR) (2003)

[4] “Cyber-physical security for the IoT society Research and Development Plan”, [https://www8.cao.go.jp/cstp/gaiyo/sip/keikaku2/3\\_10t.pdf](https://www8.cao.go.jp/cstp/gaiyo/sip/keikaku2/3_10t.pdf) (2020)

[5] TJ. Park, and SH. Hong, “Experimental Case Study of a BACnet-Based Lighting Control System”, IEEE Transaction on Automation Science and Engineering, Vol.6, No.2 (2009)

[6] K. Nakada, and H. Nakano, “Causes and Prevention Measures of Ignition and Fire from Electrical Installation”, The Journal of the Institute of Electrical Installation Engineers of Japan, Vol.29, No.8, pp.612-615 (2009)

[7] H. Ito, “BACnet System Interoperability Guideline”, The Journal of the Institute of Electrical Installation Engineers of Japan, Vol.32, No.2, pp.127-130 (2012)

[8] ST. Bushby, HM. Newman, “BACnet today”, ASHRAE journal, Vol.44, No.10, pp.10-18 (2002)

[9] SH. Hong, and S. Lee, “Design and Implementation of Fault Tolerance in the BACnet/IP Protocol”, IEEE Transactions on Industrial Electronics, Vol.57, No.11, pp.3631-3638 (2010)

[10] CA. Hollinger, “Strategies for Using BACnet/IP”, Measurement & Control Technology, Vol.5, pp.113-117 (2012)

[11] K. Tomizawa, “Object and Property”, The Journal of the Institute of Electrical Installation Engineers of Japan, Vol.32, No.2, pp.115-118 (2012)

[12] T. Toyoda, “BACnet Objects Related to Physical Access Control System”, The Journal of the Institute of Electrical Installation Engineers of Japan, Vol.35, No.10, pp.732-736 (2015)



# Process Improvement of Quantitative Progress Management Process

†Akihiro HAYASHI

†Shizuoka Institute of Science and Technology, Japan  
pixysbrain@gmail.com

**Abstract** - Process improvement has already been discussed for nearly 20 years, but the actual improvement effect has not been confirmed. In this research, we focus on the progress management process, which is most frequently used in the development life cycle. First, we define workload as the parameter that is the basis of progress management, and estimate project workload in different ways during the project planning stage. Then we define the progress management way by using workload. By defining Project Planning Stage, Progress Management Stage and Organizational Process Management Stage, divergence between estimation and actual result should be improved. When this proposed method is applied to a company that achieved CMMI level 3, an actual improvement effect is confirmed in the accuracy of workload estimation.

**Keywords:** Process Improvement, Progress Management, EVM

## 1 Introduction

Process improvement in system development has been widely discussed since the release of the CMM Ver1.1 software in 1993. The main idea was to accurately define the requirements in the upstream process and accurately estimate the scale and construction period of the project, thereby eliminating any delays in the downstream process. As a basis for this, a graph was presented showing that if modification occurred in the coding stage of the project life cycle, the modification cost would be less than half in the upstream process, but if modification occurred in the downstream process, it would increase exponentially. To prevent this type of retrogression, it is important to manage upstream processes.

In this research, focusing on the most frequently used progress management process in system development, we propose process improvement. Progress management is the process of measuring the progress of a project according to the project plan and correcting any discrepancies. The progress management process is basically implemented at progress meetings. In many companies, progress meetings are held weekly. In some industries, it is not uncommon to hold progress meetings every morning. Thus, progress meetings are held frequently, and we can expect to enhance the process through the improvement of the productivity of the progress management process by reducing rework, as well as the time and effort required for progress meetings.

As a prior study in the process improvement field, Fukuyama et al.[1] report that in process improvement using models, such

Table 1: Effort Estimation Adjustment

Top Down Approach	Conversion to Effort using the CoBRA method
Bottom Up Approach	Calculation of Effort by stacking WBS

as CMMI, there is no description of what to improve, but how to improve. Sakamoto et al. [2] propose a method for motivating process improvement by describing and analyzing the development process in a formal manner and quantitatively predicting the amount of man-hour reduction to be achieved to demonstrate specific benefits. Tanaka et al.[3] describe the organization's current process and report cases of conducting process improvement activities by motivating them via the estimation of the improvement results. Tanaka[3] reports cases of motivation by describing the organization's current processes and estimating the improvement results.

Regarding the progress management process, Okamura[4] has established a progress management methodology by using project management body of knowledge (PMBOK) in the case of IBM's giant projects. Horiguchi et al. [5] propose a management system by collecting and analyzing the access status of lecture materials as a browsing log. To make earned value management (EVM) easier to understand, Kino et al.[6] clarify the differences between traditional methods and EVM, and thereafter consider the effective implementation of EVM. Furthermore, prior research on process improvement and progress management has been reported ([7]), but no prior research has been proposed on how to manage progress consistently from project planning to project completion.

Below, in this article, we point out the lack of progress management practices and problems to be solved; in Section 3, we propose a method of project progress management for master schedule and work breakdown structure (WBS), and a method of progress management based on man-hours by EVM, while in Section 4, we discuss the results, while the conclusions are presented in Section 6.

## 2 Practice and Challenges Required for Progress Management

### 2.1 Factor to be solved

In CMMI, the most standard process management model currently being used, the equivalent of progress management

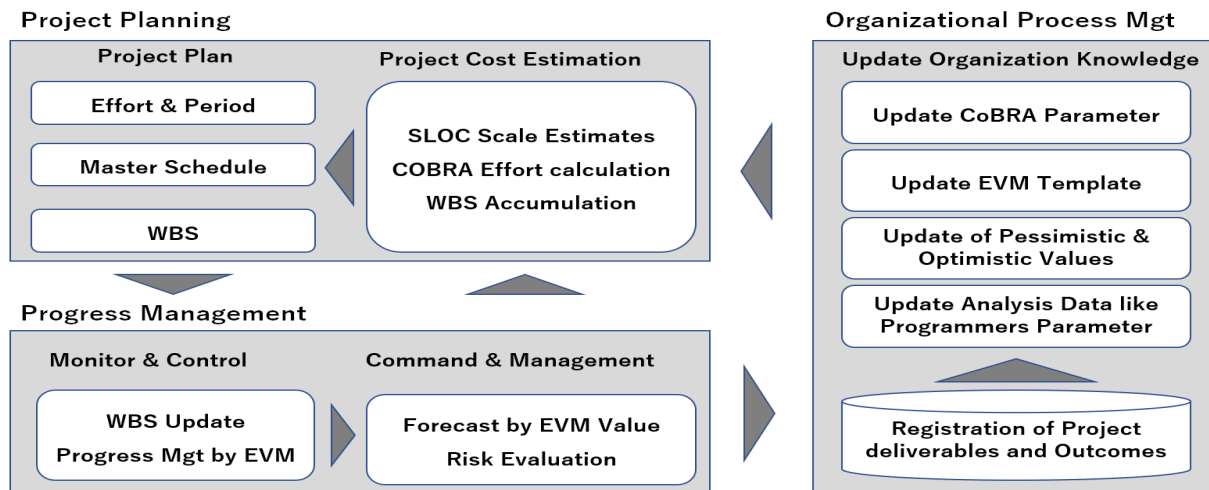


Figure 1: Relationship between project plan, progress management, organization process asset

Table 2: WBS and WBS Dictionary

WBS			WBS Dictionary					
Phase Name	WBS Code	Work Package	Author	Estimated Start Date	Estimated End Date	Estimated NOS Items	Average Effort	Estimated Effort
Design Phase	Screen Design	State transition diagram	AAA	MM/DD	MM/DD	10 items	2 Days	20 Days
...	...	...	...	...	...	...	...	...

is project monitoring and control described. This process describes the following goals and practices.

Here, SG1 (goal 1) lists management items for each progress meeting, and SG2 (goal 2) lists management items for issues and corrective actions. In other words, SG1 is described as "weekly management items" and SG2 is described as "weeks-long management items". In other words, progress management practices described in CMMI are intended for progress meetings, which are the main areas of progress management.

Notably, CMMI and other so-called best practices are written about what to do, but not how to do it. Organizations that have achieved CMMI level 3 have successfully introduced the progress management process at that point, but it is debatable whether this will have specific improvement effects on the progress management process. Adopting a best practice model to implement progress management has not established a methodology that will contribute to the overall productivity of the project, which is a challenge to be solved.

- Absence of project baselines (Factor 1)
- Feasibility has not been assessed (Factor 2)
- Lack of skills related to WBS (Factor 3)
- Lack of cost progress measures (Factor 4)

### 3 Continuous Improvement of Progress Management Process

In Section 3, to solve the problems presented in Section 2, we propose a method for continuous improvement of the progress management process shown in Figure 1. First, we present the basic policy of this research and thereafter comprehensively explain the method proposed in this research.

#### 3.1 Project Plan

##### 3.1.1 Project Effort and Construction Period

First, we define the requirements as detailed as possible to start a project, and use source lines of code (SLOC) as the scale of the new system. In this study, we do not discuss the method of scale estimation using the SLOC. Next, we use top-down and bottom-up approaches to estimate the effort.

The top-down approach refers to the CoBRA method, which converts SLOCs to effort. The CoBRA is referred to as an estimation method that visualize your "guess." For example, it quantifies your experience values, such as user communication, level of performance requirements, ambiguity of requirements, and system complexity as risks of system development. The effectiveness of the CoBRA method has been confirmed in many previous studies [8].

The bottom-up approach to man-hour estimation refers to the accumulation of man-hours required to create the work



deliverables described in the WBS.

The bottom-up approach to effort estimation refers to the accumulation of effort required to create the work deliverables described in the WBS. For an organization that introduces CMMI, a configuration management plan is prepared at the time of project planning, and the work output (including source code) and number created for each process are identified and assigned to the person in charge. If a person takes 5 days to produce a work deliverable, it takes 50 days to produce 10 work deliverables. By accumulating the effort created for each process, the overall effort of the project is determined from the bottom up.

The effort obtained by the top-down approach is the overall effort of the project. There is no perspective of when, who, what, or how to work. The effort obtained by the bottom-up approach is cumulative, concurrent tasks are not considered, and estimates are calculated from the values obtained from the top-down and bottom-up approaches to calculate the approximate effort of the project. The results are presented in Table 1.

### 3.1.2 Master Schedule

Next, a bird's-eye view of the entire project master schedule is created. The master schedule is a sheet of paper that lists all the processes, milestones, and major events from start to finish of the project—Program Evaluation and Review Technique (PERT) projection.

Once the master schedule is completed, the critical path refers to a path that cannot be accelerated owing to the dependency of the processes and tasks. By applying the organization's development phase to the critical path and visualizing the start, inter-process dependency, minimum lead time, and so on, reliable process management becomes possible.

In the master schedule, the development process is described up to the WBS code with another stage breakdown, and the required time is expressed as the length of the horizontal line of the PERT. The time required is determined by pesA, a pessimistic value that refers to the worst-case scenario. An optimistic value refers to the smoothest progress. A standard value refers to the realistic value of the project progress, and the duration is determined using the following formula and a master schedule with a quantitative basis.

$$\text{RequiredPeriod} = (\text{PessimisticValue} + 4 \times \text{MaximumPossible} + \text{OptimisticValue}) / 6$$

### 3.1.3 WBS

Once the master schedule is created, the effort for the entire project and the start and end dates are fixed, and you will know how long it takes for each development period. The WBS is created as depicted in Table 2. The left side of the Table 2 is "narrow WBS" which is the top-level content consistent with the master schedule. The "WBS Dictionary" on the right side

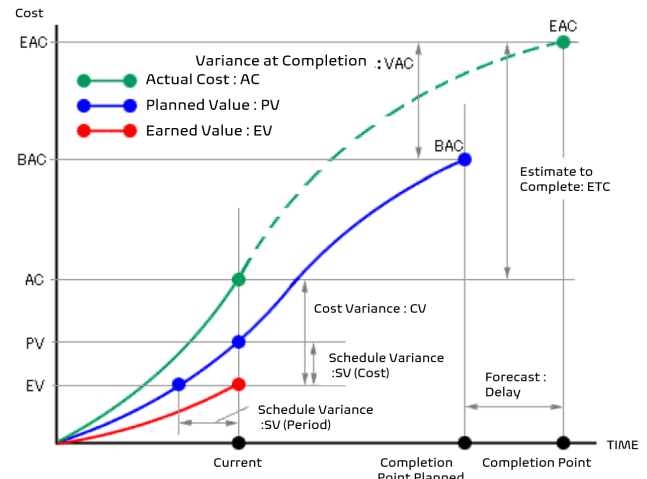


Figure 2: Concept of EVM

of Table 2 refers to the person in charge, the necessary man-hours, the start date, and the end date, and the specific tasks used for progress management.

It is difficult to do detailing the WBS at the project planning stage, so the project plan is equivalent to the master schedule. However, when the project starts, each development phase breaks down to manageable tasks. A manageable task is to break down a person's task to no longer than 5 days, start–end date, and effort.

## 3.2 Project Progress Management

### 3.2.1 Update EVM

The project progress management process is carried out in progress meetings held by project bodies. In many projects, progress meetings are held once a week, and the WBS is updated accordingly.

In updating the EVM, the cost accounting standard for activities will be "50%-50% rule". When the work begins, 50% of the estimated man-hours will be recorded as progress. After that, the remaining 50% will be recorded upon completion, not until completion. The EVM's accounting methods include "0%-100% rule," "20%-80% rule," and so on. In this study, simplicity is prioritized over strictness. Because Progress management is held once a week and activities are within 5 people's days, so even if progress is delayed, most of them are expected to be caught up the following week.

By updating the EVM, the parameters at that time are obtained. Figure 2 shows the EVM concept and the parameters used in the progress management. This graph shows the costs on the vertical axis and the time on the horizontal axis. In the planned value (PV) of this graph, the proficiency curve (S-curve) is the cost baseline. This is the planned cost of an activity based on a schedule.

The EVM basically understands PV, EV, and AC at the project site, and other values can be calculated by calcula-

Table 3: Measure of Progress

ECM Express	Changes
SV is -5 days	Report SV values instead of reports like "one week delay"
ETC is 50 man-days	ETC is reported to understand "So What"
SPI is lower than 0.9	Quantitatively reports the delay, not the "slightly late" report
VAC is 50 man-days	VAC is reported as a quantitative value of delay

tion, so the other parameters are calculated by the calculation formula built into the WBS form.

### 3.2.2 Progress Report

In previous progress meetings, there have been many subjective and ambiguous reports. This study eliminates this ambiguity and provides objective reports using EVM values as shown in Table 3 that describe the measure of progress.

This study regards man-hours as the cornerstone of project progress management. It predicts future man-hours excesses, always conscious of the cost difference of completion, variance at completion (VAC). The EVM volume is equivalent to the cost. Project progress management uses "man-day" frequently, so costs can be regarded as "man-day." It has been reported that the expression "SV is -5 man-days".

### 3.2.3 Progress analysis and risk assessment with EVM

After receiving the report from each person in charge, we conduct a progress analysis and risk assessment using the parameters of the EVM. The standard values for the progress analysis are listed in Table 4.

In the progress analysis, if the SPI or CPI reported exceeds the "standard value" in Table 5, it will be noticed as 'smokey'. If the delay continues more than 3 weeks a, or if the delay increases, it will be judged as 'fire-extinguishing' and 'fire-extinguishing team' will be introduced.

In risk assessment, a time series analysis of SPI and CPI is performed which describe in Figure 3. A graph is created with CPI on the vertical axis and SPI on the horizontal axis, and each delay is visually identified as accelerating or stopping

Table 4: Measure of Progress

Measure	Criteria	Description
SPI	<0.9	Schedule progress delayed by 10%
CPI	<0.9	Cost Progress delayed by 10%
SPI*CPI	<0.8	Development Productivity delayed by 80%
EAC/BAC	>1.1	10% cost over when completed

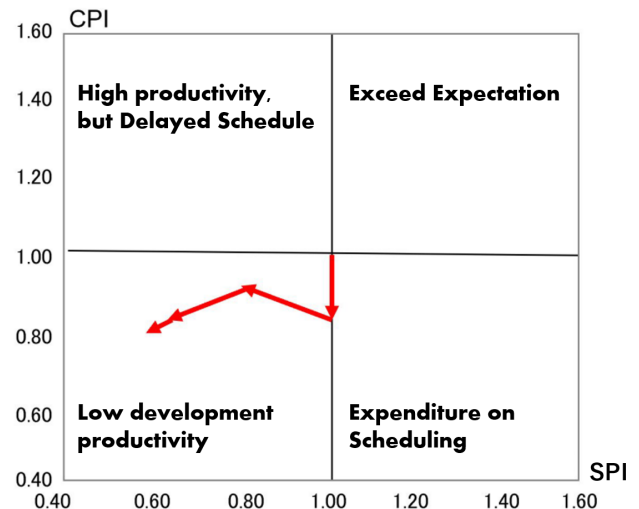


Figure 3: SPI/CPI Time Series Analysis Example

falling. If CPI decreases at the same time as SPI, development productivity ( $SPI \times CPI$ ) will decrease, and man-hours will exceed expectations significantly. Alternatively, if you know that you are sacrificing CV to maintain SV from a certain point in time, you can estimate that you are compensating for it by working overtime or taking time off if you do not put in additional factors.

Finally, we will use EAC/BAC to determine the overall status of the project; if the EAC/VAC value exceeds the threshold (e.g., 110%), it means 10% or more of the cost overruns of the plan at the end of the project. These escalate to the senior management who make decisions such as acceptable project cost deficits and push customers for extra staff to ensure quality.

## 3.3 Organization Process Asset

### 3.3.1 Project Work Deliverables Registration

After completing the project, all the work products and process assets of the project will be registered in the organizational process asset database.

Generally, the means to increase productivity include increasing the mechanization rate, increasing concurrency, improving utilization rate, and increasing reuse rate. In the case of manufacturing, such as system development projects, improvements in reuse will greatly contribute to productivity improvement.

In addition, activities such as project management can inevitably lead to trial and error. By reusing the process assets of successful projects, one can expect to save the effort of trial and error.

Organizations that use CMMI plan the work output for each development phase in the configuration management plan. Keep a record of the work output consistent with this plan. In principle, all the work output of the project should be registered in the organization process asset database and reused in the next project.

### 3.3.2 Register Lessons Learned

In this proposed method, decisions are made in the following parts of the project using intuition and subjective judgments.

- In estimating man-hours, we combine the top-down approach and the bottom-up approach to make estimates. At this time, we do not simply take intermediate values, but make decisions based on the knowledge of experienced people.
- Using the CoBRA method, we use parameters that express the knowledge of experienced people as quantitative values.
- In the PERT projection of the master schedule, pessimistic and optimistic values are calculated using rules of thumb when defining the period of the development phase.
- Progress analysis was conducted using EVM values at the progress meeting, and thresholds for values such as SPI/SPI, VAC/EAC were determined.
- Based on the project period, we use a rule of thumb to determine how long the delay can be recovered.

Once the project is completed, we will verify whether such a subjective judgment is correct and put it in the process asset database as a lesson learned. These lessons are not something that can be used as is, so we aim for spiral-up by turning this cycle many times.

If you do not have the experience, it is not good to think that you can ask the experienced even if you are in the same organization. Since the rules of personal experience are tacit knowledge, implicit knowledge is likely to be lost for retirement or other reasons. You must quickly formalize it and incorporate it into organizational knowledge.

## 4 Application Result

The results of the application of the proposed method of this research at Company C, which appeared in the case analysis in Section 2, is explained herein. Company C is developing embedded software for measuring instruments. In manufacturing measuring instruments, derivatives such as enhancements are often used. Because so called Function Point The FP method is difficult to apply in the development of embedded systems with precision instruments.

When the company applied this proposal method, SLOC estimated the scale of modification of the previous derivative development of the same model and the area to be modified this time. Next, considering the frequency of customer demand changes, difficulty of neck technology, and frequency of interface changes, the CoBRA method is used for conversion into man-hours. Next, we created a WBS for this development by reusing the WBS from the previous derivative development

project and estimated the man-hours by accumulating the described man-hours. Finally, we calculated the man-hours of this project by combining the two.

After the project is completed, the accuracy of the estimate is measured using the VAC value. The following graph shows the change in estimation error for about three years after the introduction of this proposed method. In the first year, the difference between estimates and VAC performance was more than 20%. Three years ago, it improved by 3%. This company has adopted only this proposal method, so the improvement effect can be attributed to this proposal method.

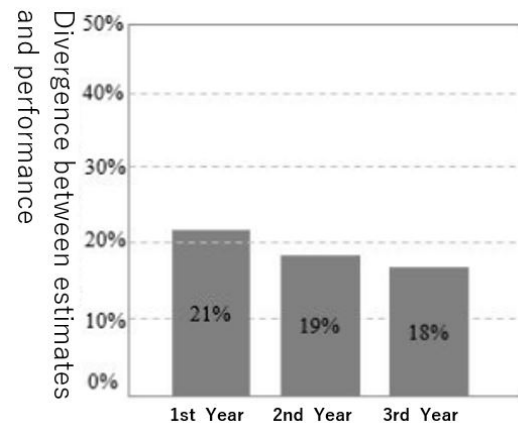


Figure 4: Changes in VAC Ratio to Total Effort

## 5 Conclusion

In this study, we focused on the fact that process improvements in system development are targeted at "large PDCA" for the entire life cycle, and proposed improvements with "small PDCA" covering progress management; best practices, such as CMMI, include what to do, but not how to do it, so we proposed man-hour-based practices, such as estimating man-hours at the project planning stage, meeting progress at the process management stage, and registering in process assets after the project is completed.

By using this proposed method, we eliminate subjective and ambiguous progress reports, such as "delayed but good," that have been held at previous progress meetings, and have been able to quantify the risk assessment using EVMs, so we can see how much the cost will be overrun at the end of the project.

However, this study did not evaluate the EVM threshold. For example, the SPI/CPI threshold started at 0.9 because excellent companies such as IBM and Unisys set the threshold at 0.9. However, the pros and cons of small and medium-sized enterprises imitating the practices of large enterprises such as IBM have been repeatedly discussed, and if  $SPI < 0.9$ , the risk of becoming obsolete is indicated. It is reasonable to spiral up by accumulating the lessons learned. Establishing the correct thresholds for each development site will be a challenge in the future.

## REFERENCES

- [1] S. Fukuyama, S. Miyamura, H. Takagi, and R. Tanaka, "A Software Process Improvement Support System: SPIS," IEICE Trans. Inf. & Syst., Vol. E83-D, No. 4, pp. 747-756 (2000).
- [2] K. Sakamoto, et al., "An Improvement of Software Process Based on Benefit Estimation," IEICE, D-I, Vol. J83-D-I, No. 7, pp. 740-748 (2000).
- [3] T. Tanaka, K. Sakamoto, S. Kusumoto, K. Matsumoto, and T. Kikuno : "Improvement of Software Process by Process Description and Benefit Estimation", 17th Int'l Conference On Software Engineering, pp.123-132 (1995)
- [4] S. Okamura, Progress Management in Project - Master Schedule, Earned Value Management and Project Management System, Nikkei BP, (2010).
- [5] S. Horiguchi, Progress Management Metrics for Programming Education of HTML-based Learning Material, IPSJ Journal, Vol. 53, No. 1, pp. 61-71, (2012).
- [6] Y. Kino, Progress management and EVM based on volume of deliverables, Project Management Institution, Vol. 5, No. 3, pp. 11-15, (2003).
- [7] H. W. Tuan, C. Y. Liu, and C. M. Chen, "Using ABC Model for Software Process Improvement: A Balanced Perspective," Proceedings of the 39th Annual Hawaii International Conference on System Sciences (HICSS'06) (2006).
- [8] Introduction to CoBRA Method, CoBRA Laboratories, Ohmsha, (2011).
- [9] Y. Mizukami, M. Ida, "Improvement of Workload Estimates by using project characteristic based on CoBRA Method for Software Development Project", Journal of Japan Society of Directories, Vol 11, pp. 36-45, (2013).
- [10] Dai Sakai, The Points of Building Estimating Model Using CoBRA Method, SEC journal, Vol. 7 No. 3, Oct. 201