IWIN2011

International Workshop on INformatics

Proceedings of
International Workshop on Informatics

September 16-21, 2011
Ca' Foscari University, Venice, Italy

IWIN2011

International Workshop on INformatics

Proceedings of
International Workshop on Informatics

September 16-21, 2011
Ca' Foscari University, Venice, Italy

Table of Contents

Keynote Speech 1 (13:30-14:30, Sept. 19)

Session 2: Business models and Security (Chair: Takuya Yoshihiro) (14:50-16:30, Sept. 19)

Session 3: Network Routing and Multimedia (Chair: Hideyuki Takahashi) (9:50-11:30, Sept. 20)

Session 4: Computer Assisted Work (Chair: Yoh Shiraishi) (13:30-14:30, Sept. 20)

Session 5: Mobile/Pervasive Computing (Chair: Yoshitaka Nakamura) (14:50-16:30, Sept. 20)

Session 6: Social Studies (Chair: Tomoya Kitani) (9:30-10:20, Sept. 21)

Keynote Speech 2 (10:20-11:20, Sept. 21)

A Message from the General Chair

Welcome all of you to the fifth International Workshop on Informatics (IWIN2011) once again in Italy this year. The first workshop was held in Naples, Italy in 2007. The workshop was the first event for the Informatics Society. We had the workshop in Austria in 2008, Hawaii, U.S.A. in 2009, Scotland and U.K. in 2010. This year, IWIN2011 is held in a historical place in Venice, Ca' Foscari.

The Keynote talks are the highlight of the workshop. This year we have two keynote speakers. The first speaker, Prof. Peter T. Kirstein from University College London, is one of the pioneer researchers in the networking area and has been working over forty years. He shall present the recent research issues in this area. He shall talk on IPv6.

The second speaker, Dr. Kazuhiko Kushima from Nippon Telegraph and Telephone Corporation (NTT) has been working on various aspects of business and research with NTT. He shall talk on current services and research tends.

We have carefully selected 23 papers as well as another four papers at the pre-workshop. They cover a wide range of areas in informatics such as networking including recent issues on home and ubiquitous networks as well as mobile and pervasive networking , multimedia, security, computer assisted work, and social studies. The pre-workshop which includes the papers in the evaluation of research in information systems, video sharing, life log and cloud computing.

I would be grateful for the arrangement made by Prof. Agostino Cortesi, our colleague of the program committee as well as the local chair, who made it possible for us to hold the workshop in such a lovely venue at Ca' Foscari.

I thank Dr. Tomoya Kitani from Shizuoka University who has done most of the work to prepare the workshop. Without him the workshop was not possible.

Sincere thanks to Tadanori Mizuno who is the head of Informatics Society for his continuous encouragement towards us to hold the workshop.

Last but not least, I would like to thank all the participants and contributors who have made the workshop possible.

We are looking forward to seeing you all in Venice.

Yuko Murayama
General Chair
The International Workshop on Informatics

Organizing Committee

General Chair

Yuko Murayama (Iwate Prefectural University, Japan)

Local Arrangement Chair

Agostino Cortesi (Ca' Foscari University, Venezia, Italy)

Steering Committee

Toru Hasegawa (KDDI R&D Laboratories, Japan)

Teruo Higashino (Osaka University, Japan)

Tadanori Mizuno (Shizuoka University, Japan)

Jun Munemori (Wakayama University, Japan)

Yuko Murayama (Iwate Prefectural University, Japan)

Ken-ichi Okada (Keio University, Japan)

Norio Shiratori (Tohoku University, Japan)

Osamu Takahashi (Future University-Hakodate, Japan)

Program Committee

Agostino Cortesi (Ca' Foscari University, Venezia, Italy)

Kouji Yoshida (Shonan Institute of Technology, Japan)

Gen Kitagata (Tohoku University, Japan)

Hiroshi Mineno (Shizuoka University, Japan)

Hirozumi Yamaguchi (Osaka University, Japan)

Teruyuki Hasegawa (KDDI R&D Laboratories, Japan)

Tomoya Takenaka (Shizuoka University, Japan)

Teruhisa Ichikawa (Shizuoka University, Japan)

Yuko Murayama (Iwate Prefectural University, Japan)

Jun Sawamoto (Iwate Prefectural University, Japan)

Masahiko Ishino (Fukui University of Technology, Japan)

Yoshia Saito (Iwate Prefectural University, Japan)

Fumiaki Sato (Toho University, Japan)

Takaya Yuizono (Japan Advanced Institute of Science and Technology, Japan)

Taizo Miyaji (Tokai University, Japan)

Tomoya Kitani (Shizuoka University, Japan)

Hidekazu Shiozawa (Tamagawa University, Japan)

Tomoo Inoue (University of Tsukuba , Japan)

Hirosato Tsuji (Mitsubishi Electric Corporation, Japan)

Kozo Okano (Osaka University, Japan)

Toshihiro Shikama (Fukui University of Technology, Japan)

Shohei Yokoyama (Shizuoka University, Japan)

Hiroshi Shigeno (Keio University, Japan)

Yoh Shiraishi (Future University-Hakodate, Japan)

Shinichiro Mori (Fujitsu Laboratories, Japan)

Takuya Yoshihiro (Wakayama University, Japan)

Tsukasa Kudo (Shizuoka Institute of Science and Technology, Japan)

Masakatsu Nishigaki (Shizuoka University, Japan)

Satoru Fujii (Matsue College of Technology, Japan)

Yuji Wada (Tokyo Denki University, Japan)

Jun Munemori (Wakayama University, Japan)

Preworkshop
(Chair: Tomoya Kitani)

A study of Effectiveness Evaluation in Information Systems

Teruhisa Ichikawa[*], Kiminobu Kodama[**] and Tomoya Kitani[***]

[*]Faculty of Informatics, Shizuoka University, Japan

[**]Information Systems Institute, Ltd., Japan

[***]Division of Global Research Readers, Shizuoka University, Japan

Abstract - In the information systems field, the composition of a report is difficult. There are few contributions, and only a few of these are adopted. The evaluation of the effectiveness of research is the most important factor to write a paper properly in preparing a proper document. The purpose of this research is to propose the model evaluation technique peculiar to determining the effectiveness in information systems and recommending new criteria. We also touch briefly on the evaluation of the object system, the improvement in academic value, and suggested "trouble-shooting" for remedial measures.

The effectiveness evaluation subcommittee of the Information Processing Society of Japan (IPSJ) was founded last September. The subcommittee promotes this research. Current activities give priority to the analysis of system papers adopted in the past. In this way the technique and criteria for evaluation are clarified.

This paper is a summary of problems raised by the results of current analysis procedures. The first problem is that the presentation of statistical information in charts and graphs is not appropriately and consistently applied. The second problem is the difficulty of applying quantitative research to an information system which has an individual, singular character and which is unique. The first problem may be resolved by establishing a protocol of introducing the method of generating the statistical information with explanatory notes for each evaluation pattern. A protocol establishing the context and application experience of the qualitative study will be effective for the second problem.

Keywords: information system, effectiveness evaluation, evaluation model, criterion

1. INTRODUCTION

In the information systems field, the composition of a report is difficult. There are few contributions, and only a few of these are adopted. The evaluation of the effectiveness of research is the most important factor to write a paper properly in preparing a proper document. The purpose of this research is to propose the model evaluation technique peculiar to determining the effectiveness in information systems and recommending new criteria. We also touch briefly on the evaluation of the object system, the improvement in academic value, and suggested "trouble-shooting" for remedial measures.

If a new contributor makes use of the present research to successfully analyze his/her own work, one can easily understand the academic value of our proposals. Moreover, ideas for improving the system of one's own company can be obtained by utilizing our analytic system within the enterprise itself.

1.1 background of research

Authors are members of the editorial board for IS research papers. Few IS papers are submitted and the rate of selection is low. We are attempting to improve this situation with various approaches.

(1) Discussions concerning the writing of IS research papers are conducted frequently. The formats of large number of IS research papers were analyzed and the results compiled in "Proposal and Submission Criteria for IS Research Articles" [1]. A campaign to disseminate this educational material is being conducted and the above mentioned "Proposal" is becoming standard in the field.

(2) Every year a research paper writing workshop is conducted by volunteer IS research reviewers. The papers are reviewed with writing points, evaluation standards, and the inclusion of samples of rejected works. We aim for a higher acceptance rate.

(3) Each year we produce a special issue devoted to IS research. We invite proposals, screen, and then publish them. It reports on the result in IPSJ Bulletin [2][4] and the proceedings of IS conferences within IPSJ [3][5][6][7]. Improvement is thus registered.

However, there has not been an improvement in the number of submissions and the inspection turnover.

According to the regulations covering system development papers submitted to the Institute of Electronics, Information and Communication Engineers, contributors are required to present "New use concerns the latest elemental technology and that the effectiveness be proven objectively."

According to the criteria for selection for the IPSJ, the original research must be novel. Novelty is described as never before proposed, but may be a novel combination of previously utilized ideas or suggestions for new uses. Effectiveness requires that we evaluate it from the viewpoint that the utility of the proposal technique is shown by the performance assessment etc. or that technical effectiveness is objectively confirmed by the work and the product (software and hardware, etc.), for it to be commercialized or available to the public.

Nagata [1] stated that the relations of humans with the enterprise and society are complex and so it is difficult to evaluate the novelty, efficacy, and authenticity of research concerning information systems.

For instance, even if a new inventory control system is designed and the inventory cost is reduced, it does not necessarily prove the efficacy of the new system. The business environment changes hourly. The product line, parts, and the

form of dealings, etc. are changing, too. One cannot judge whether reduction in inventory cost is the direct effect of a new system or whether other factors must be considered. There is a method of demonstrating the dominance of system compared with a similar inventory control system. However, because the product and the environment change, the comparison is not significant.

Even if a method is useful for the business of one company, if it cannot be applied to other similar businesses, it cannot be said to be an effective method. There may even be a lot corresponding elements in the businesses, but when one begins to design the method for application to the other business, the application does not proceed.

The IS feature is singular; therefore, the analytical approach does not have easy application. It is difficult to demonstrate effectiveness by reference to a past standard.

Kaminuma [3] described IS work as follows: IS study evolves only after the theory and a technological application are played out in the practice field. The IS research object is an interdisciplinary mapping of the use environment (person, organization, and society) etc. that surrounds the practice and application technologies. Because of the uniqueness of each context, many IS papers were case studies. Novelty of a new system cannot be appreciated without an adequate survey of previous operations and comparison with a similar system. There are indispensable to showing the novelty.

In IS research, it is necessary to perform a logical evaluation to demonstrate utility (effectiveness) to the reader. Many IS submissions lacking this requirement failed in the selection process.

It is challenging to structure research in the IS field that does not lend itself to codification in research publications. Nevertheless, some IS submissions succeed in getting over the wall of inspection. Various devices are employed to prove the effectiveness of IS research papers. By selecting and recording these devices and a common pattern emerges. This pattern makes a foothold, to be correctly drawn and modeled. It was thought that this work one strategy that climbs the wall of the obstruction to structuring a research paper easily.

1.2 Objective of the research

This study aims at modeling an evaluation technique for the peculiar effectiveness of IS. Specifically, we analyze selected IS research papers which proved effectiveness in the past and categorize the factor to be picked up by the method. The categories which were determined through the analysis of the collected past papers are as follows:

1) Object of systematization
2) Environment of the system
3) Purpose of systematization
4) Aspect of efficacy evaluation
5) Technique of evaluation

These factors are systematized and mod-

eled. In addition, we aim to propose a new criterion and an evaluation method suitable to the peculiar IS values.

This research was promoted by the Information System Society of IPSJ, and the fiscal year target is set up in "Effectiveness evaluation technique subcommittee" for the conference in September, 2010.

The subcommittee is executing the following schedule:

· Fiscal year 2010. Collect IS papers selected and recorded in the proceeding of academic societies connected to the key information. After analyzing the various method of evaluation, make a database allotting categories.

· Fiscal year 2011. Examine a more appropriate questionnaire form, the statistical work, and the applicability of the qualitative research method such as GTA. In addition, make use of the accumulated database, evaluation systematization and modeling and evaluate the adaptability of these instruments. A new criterion and method are proposed based on the effectiveness evaluation model so constructed.

· Fiscal year 2012. Review the structure of the database. Systematization the collection of case studies and IS research papers for analysis. As a result, user's convenience is improved. Moreover, the effectiveness evaluation model, the standard questionnaire form, and the statistical analysis example are collected to the catalog. It prints on the paper medium, it distributes in the society and the workshop, and the educational campaign is done.

A lot of members are participating from the university and the enterprise in "Effectiveness evaluation technique subcommittee". The member analyzes the paper allotting it. It has a session at the frequency once every two months, it introduces research results each other, and the discussion is deepened (Ref. Figure 1).
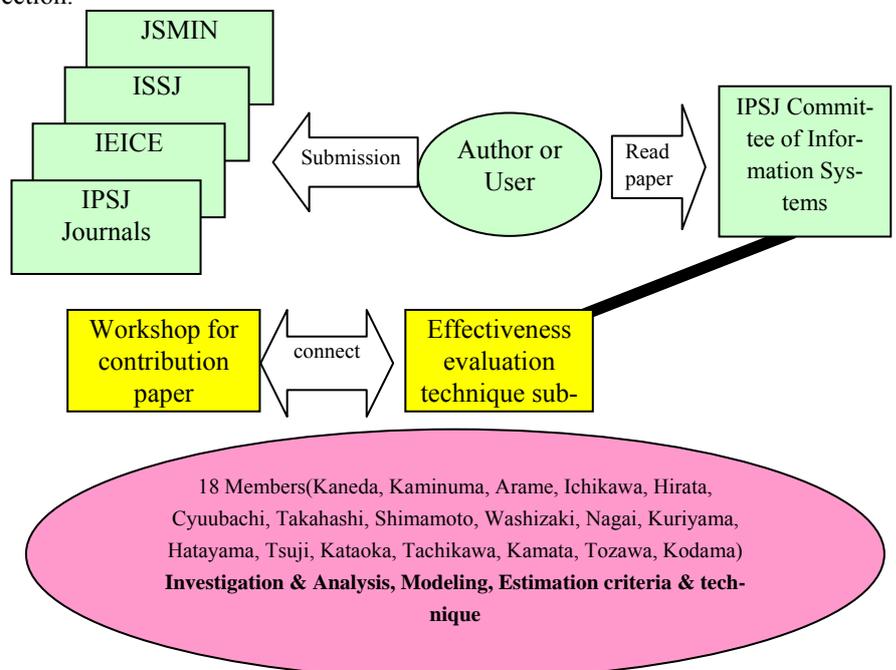
Figure 1. Research organization of effectiveness evaluation technique

2. SURVEY ANALYSIS OF PAPERS

We collected IS related paper from journal of the main academic society related to the information. And, we analyzed the way of the evaluation used with those papers. The corresponding one was chosen from the journal 1990 to 2010 as a present analysis target (Refer to the analysis paper list).

The Japanese paper: It retrieved from the journal of IPSJ (Information Processing Society of Japan) by key word "Information system" and "Evaluation", and 27 corresponding papers were chosen. For instance, the following paper corresponds.

· Models and Techniques for Measuring Customer Satisfaction so as to Evaluate Information Systems
· A Case Study for Information Systems Assessment : Management Roles of Information Systems in Overseas Subsidiaries

An English paper: It retrieved from CACM (Communication of the Association for Computing Machinery) and JAIS (Journal of the Association for Information Systems) by key word "Evaluation of IS", and 11 corresponding papers were chosen. For instance, the following paper corresponds.

· A Cost-Benefit Evaluation Method for Web-Based

38 selected papers were analyzed by the subcommittee members. It allocated it at random so that two people a volume might take charge. Because two people independently evaluated the same paper, it considered it not to be one-sided of the opinion. The analysis result was filled in the evaluation seat shown in Table 1.

Table 1. Evaluation seat of analysis

Paper title	Author	Publishing	Abstract	Objective of the system	
Estimation technique					Estimation result
Compare	measurement	environment	method	others	
Opinion of analyst					
Estimation of estimation technique	Problem & improvement	notice	Important reference		

It analyzed it concerning 35 papers except three papers judged not to have treated the information system obviously in 38 papers. We met it once every two months. We discussed a preferable evaluation method from the analysis result. The analysis item was reviewed if necessary.

The questionnaire is used well to evaluate the effectiveness of the information system. There were essentially a lot of questions of insignificant question, for instance, "Was this system useful?" There were a lot of papers that did a statistical analysis not significant and had lowered the research quality. Therefore, illustrating a more appropriate questionnaire form and the statistical work also examined it concurrently. However, this research is the one that how to treat what should be and the acquired data of the research is discussed. Because it is not the one to study writing and a statistical technique of the paper, it examines it enough.

The following have been understood from the way analysis shown in Table 2.

1) For comparison
   The comparison shows the way to prove the domination to be evaluated. It is necessary to show the validity of the compared object in "Compare it before and behind", "Class comparison", and "Standard value comparison". It is necessary to show the validity of the hypothesis in "Temporary housing verification". It is necessary to show the design adequacy in "Operation verification". The system will be mounted as a means to prove utility from each standpoint. However, many of indices were the ratings by the questionnaire survey, and there were a lot of paper to lack objectivity.

2) For measurement value
   The measurement value shows it tried to evaluate it what measuring. In general, the authors like physical value that possible to measure automatically and easy to treat as objective data. However, there was a lot of paper that treated the amount of psychology.

3) For measurement environment
   It is preferable to measure it in a real environment. However, there were a lot of papers pseudo measured from the reason for the influence on the business etc.

4) For measurement method
   The method of collecting the measurement values was analyzed. However, there were various methods and it was difficult to classify it clearly. The questionnaire method, the automatic measurement method, the observation, and the interview, etc. almost had the same number. There were papers with uncertain measurement method, too.

5) Others
   It is necessary to show how to analyze the measurement value and to have derived the conclusion clearly. However, the paper that was not able to trace whether to use and to analyze an appropriate method for analyzing was most. An analytical process was presented politely and plainly, and theses that were able to trace the validity of the method for analyzing were only two papers.

Table 2. Summary table of analysis result

class	technique	Number of correspond
Compare	Comparison of the before and after	11
	Comparison of similar system	7
	Comparison with managerial index	6
	Comparison between standard values and actual values	2
	no comparison	9
Measurement	measurement of physical value (time, data volume)	15
	measurement of psychological value (sense, impression)	13
	Judgment of quality and presence	5
	The qualitative data is filled in.	5
environment	Measurement on the use site	18
	Measurement on the simulation environment such as laboratories.	10
	Measurement by simulation in the computer.	4
	Theoretical description	3
method	Questionnaire	8
	Automatic measurement	7
	Observation	6
	Interview	6
	Unknown	8
others	It can be traced that an appropriate method for analyzing is used.	2
	not possible to trace	33

3. EVALUATION METHOD OF RECOMMENDED PAPERS

The opinion of the assayer in the evaluation seat includes many things. Therefore, it is difficult to classify. However, it has been understood that there are various devices to make the efficacy evaluation better. Especially, the way of the following two papers serves as a reference.
(1) Evaluation case by questionnaire

The first paper is " Models and Techniques for Measuring Customer Satisfaction so as to Evaluate Information Systems" [8].

In this paper, the index of new customer satisfaction measurement is defined after the previous work is summarized, the questionnaire survey of two actual information systems is done, and the satisfaction rating index is evaluated. The satisfaction rating is evaluated by using the rating scale value to eight factors. It is six kinds of quality properties (JIS X0129) and two kinds of value characteristics (effectiveness and effect). The weight parameter of eight factors has been decided by using the AHP method (hierar-

chical decision making method). The questionnaire of the paired comparison is totaled according to range of customers. It is indicated that the numerical result of the overall customer satisfaction measurement is a value close to the respondent's actually feeling.

In the subcommittee meeting, the following opinions were shown.
・It is realistic, and done the survey well. It evaluates with plural real systems.
・It is possible to refer if it limits within the range of customer satisfaction measurement.
・The point where it challenges the quantification of the satisfaction rating to the information system in the information system management section is appreciable.
・A part of the questionnaire actually done is published in the appendix, and the point that should be learnt in the evaluation approach of the questionnaire type.
(2) Evaluation case by multivariate analysis

Another paper is "Measuring Clarity of Compute" [9]. In this paper, the basic index that composes comprehensible of the computer manual is taken up in the hypothesis. The factor of a possible automatic measuring system is assumed to be a basic index by the morphological analysis etc. It questions about the 1st to rate "Clarity". The factorial analysis is done to the result of the questionnaire, and four factors (conciseness, legibly, understandable, friendly) are extracted. Next, the measurement model of "Clarity" is composed by using these factors.

In the questionnaire of the 2nd, those who rate it are divided into two crowds (a technical writer and the postgraduate of the engineering system).

A measurement model different according to the multiple regression analysis is derived from the rating result. It is shown to be able to measure "Clarity" mechanically by applying this model at each making stage of the manual.

In the subcommittee meeting, the following opinions were shown.
・The hypothesis is set, the conclusion is derived in statistics, and elaborate logical development that evaluates the validity in statistics serves as a model.
・The presentation of various statistical tables to the model deriving from the factorial analysis to the multiple regression analysis is a point that should be learnt.

Two papers have treated the amount of psychology (customer satisfaction measurement and plainly). The feature is logical to develop elaborate. It is necessary to examine how to discuss effectiveness to evaluate the information system quantitatively carefully. For instance, if effectiveness is shown by "merit", it is necessary to define the index to measure "merit" clearly. The index value is measured, and it statistics analyzes and it is necessary to prove the significance by showing the result of doing.

4. CONCLUSIONS

We have reported on the result of the literature research of the efficacy evaluation of the IS. In the IS related papers, the physical value at the time etc. measured in the condition similar to a real environment and the psychological elements by the questionnaire were used. The method of comparing

the measured values and showing effectiveness was adopted. However, there few papers that were able to prove effectiveness objectively.

There were only two useful reference papers according to effectiveness evaluation method of the system paper that had been obtained by the analysis during about half a year. However, the two papers do not provide a theme for a typical information system. The difficulty of the application of the quantitative evaluation method to the system paper was clarified again. The need for examining "Qualitative evaluation method" is high.

As for the research in the future, we will undertake the following three points.
1) The evaluation at design level method is examined about the effectiveness of the information system.
2) The model and the statistical analysis method of the survey slip for the questionnaire survey are maintained as a case collection.
3) The possible application to the information system of the qualitative study (for instance, grounded theory approach and action research) based on a small number of opinions is examined.

The theme of IS is socially useful. However, it remains in the case study. The meaning of arranging the analysis result as a case collection is significant. Not only the researcher but also the business person can be expected to utilize it as an index that objectively evaluates the information system that constructs it.

If the author is understood, the academic value rises. In the enterprise, the hint for considering remedial measures is obtained by objectively evaluating the effectiveness of the system. The value of a qualitative improvement of the information system that is attempted, here is substantial.

REFERNCES

[1] Nagata, M., A Proposal for Writing Reviewing Papers on Information Systems, IPSJ, 2001-IS-77 (2001).
[2] Kaminuma, Y., Editor's Message to Special Issue on Information Systems, IPSJ, 46 (3), (2005).
[3] Kaminuma, Y., Summary report of journal "Special Issues on Information Systems" and further expectation, IPSJ, 2005-IS-91 (2005).
[4] Kaneda, S., Summary Report of "Special Issues on Information Systems for New Application Domains", IPSJ, 47 (3), (2006).
[5] Kaneda, S., Summary Report of "Special Issues on Information Systems for New Application Domains", IPSJ, 2006-IS-27 (2006).
[6] Tsuji, S., Summary Report of the "Special Issue on Information Systems for Bases of Information Society", IPSJ, 2007-IS-25 (2007).
[7] Abe, A., Summary Report of the "Special Issues on Information Systems for Solutions of Social Problems", IPSJ, 2008-IS-16 (2008).
[8] Chikara, T. and Fujino, K., Models and Techniques for Measuring Customer Satisfaction so as to Evaluate Information Systems, IPSJ, 38 (4), pp. 891-903 (1997).
[9] Takahashi, Z. and Ushijima, K., Measuring Clarity of Compute Manuals, IPSJ, 32 (4), pp. 460-469 (1991).

APPENDIX: PAPERS LIST FOR ANALYSIS

Transactions of IPSJ（evaluation of IS）

1) Chikara, T. and Fujino, K., Models and Techniques for Measuring Customer Satisfaction so as to Evaluate Information Systems, IPSJ, 38 (4), pp. 891-903 (1997).

2) Miyake, T., Takenaka, I., Furuyama, T. and Nakagawa, Y., Software Quality Control Method Based on Software Quality Metrics and Its Effectiveness, IPSJ, 36 (1), pp. 105-118 (1995).

3) Miyazaki, T., Yamada, M., Shimajiri, Y., Seki, H. and Itoh, H., A String State Representing Method for Knitting and Implementing a Stitch Symbol Generating System, IPSJ, 36 (11), pp. 2736-2740 (1995).

4) Tomiyasu, H., Wakamatsu, K. and Arakawa, H., Construction of a Tablet Identification System using Parameter Design Method, IPSJ, 35 (12), pp. 2694-2702 (1994).

5) Hatsusegawa, S., A Case Study for Information Systems Assessment: Management Roles of Information Systems in Overseas Subsidiaries, IPSJ, 32 (12), pp. 1506-1514 (1991).

6) Itoh, K., Tabata, K. and Ohno, Y., System Description and Evaluation System: SDES, IPSJ, 20 (4), pp. 355-362 (1979).

7) Kawabata, R., Itoh, K. and Kumagai, S., Prototyping Based on Domain Model Considering Generic Task, IPSJ, 41 (9), pp. 2555-2566 (2000).

8) Takahasi, Z. and Ushijima, K., Measuring Clarity of Compute Manuals, IPSJ, 32 (4), pp. 460-469 (1991).

9) Amano, J. and Rikiso, Y., A User-friendly Multimedia Book Authoring System (Management of Information, Data and Knowledge) (<Special Issue> Information Systems), IPSJ, 46 (3), pp. 715-727 (2005).

10) Yano, K., Kawakami, K. and Honma, K., Evaluation of the Countermeasure for Flaming Problem on the Regional Network Community Service (Regional Information Systems) (<Special Issue> Information Systems), IPSJ, 46 (3), pp. 765-771 (2005).

11) Miyazaki, K., Ida, M., Yoshikane, F., Nozawa, T. and Kita, H., Development of a Course Classification Support System for the Awarding of Degrees Using Syllabus Data (Information Systems for Organizational Activities) (<Special Issue>Information Systems), IPSJ, 46 (3), pp. 782-791 (2005).

12) Kawamura, Y., A Study on the Product Development Scenario-Planning System Based on the Communication Semantic Chain Model (Information Systems for Society and Humans, <Special Issue> Information Systems for New Application Domains), IPSJ, 47 (3), pp. 686-700 (2006).

13) Egi, H., Shigeno, H. and Okada, K., A System for Supporting Face-to-face Collaboration Based on Dynamic Distance Zone Awareness (Information Systems for Society and Humans, <Special Issue> Information Systems for New Application Domains), IPSJ, 47 (3), pp. 711-721 (2006).

14) Sikiguchi, T. and Kato, H., Proposal and Evaluation of Video-based Privacy Assuring System Based on the Relationship between Observers and Subjects (Development and Operation of Information Systems), IPSJ, 47 (8), pp. 2660-2668 (2006).

15) Nitta, H., Takeda, H., Hochin, T., Koshimura, S. and Matsushita, T., Traceability Support System for Agricultural Products Utilizing Cell-phone (Information Systems for Society and Humans <Special Issue> Information Systems for Bases of Information Society), IPSJ, 48 (3), pp. 1010-1019 (2007).

16) Umeda, M., Katamine, K., Nagasawa, I., Hashimoto, M. and Takata, O., The Design and Implementation of Knowledge Processing Server for Enterprise Information Systems (Knowledge Processing), IPSJ, 48 (5), pp. 1965-1979 (2007).

17) Matsuzawa, Y. and Oiwa, H., A Result of Trial Education for Software Engineers through University-Industry Collaboration and Project-based Learning (Education for Computing Professionals, <Special Issue> Computers in Education: Philosophy, Theory, and Practice), IPSJ, 48 (8), pp. 2767-2780 (2007).

18) Maita, M., Abe, A., Kano, T., Kato, M. and Oshida, Y., Development and Social Experiment of a UD-based Tourist Information System by Mobile Phone and Active RFID (Ubiquitous Computing, <Special Issue> Human-Centered Universal/Ubiquitous Network Services), IPSJ, 49 (1), pp. 45-57 (2008).

19) Tsuchiya, R., Sugiyama, Y., Yamauchi, K., Fujinami, K., Arisawa, R. and Nakagawa, T., Route Choice Support System for Passengers during Disruption of Train Operations and Its Acceptability Evaluation (Information Systems for Society and Humans, <Special Issue> Information Systems for Solutions of Social Problems), IPSJ, 9 (2), pp. 868-880 (2008).

20) Inoue, A. and Kaneda, S., A PBL Approach using Real World Application Development between University and Local Government (Education for Information Systems, <Special Issue> Information Systems for Solutions of Social Problems), IPSJ, 49 (2), pp. 930-943 (2008).

IPSJ SIG Notes (evaluation of IS)

21) Kuriyama, S., Takeno, T. and Sugawara, M., A Consideration about the Evaluation of Efficiency on Investment of Information Systems, IPSJ SIG Notes, 2001 (62), pp. 31-38 (2001).

22) Kurihara, J., Nagahashi, K., Asada, T., Ishikawa, T. and Koizumi, H., Verification and Evaluation of an EC model by Web service based on SOAP/WSDL/UDDI, IPSJ SIG Notes, 2002 (30), pp. 53-60 (2002).

23) Ichikawa, H., Fukuoka, H., Oshida, Y., Kano, T. and Abe, A., Development and Evaluation of UD Tourist Information System Using Mobile Phone to Heritage Park, IPSJ SIG Notes, 2010-GN-75 (18), pp. 1-6 (2010).

24) Kaneda, S., An Evaluation Approach for Prototypes of Application Systems, IPSJ SIG Notes, 2008 (52), pp. 7-14 (2008).

25) Namba, Y., A Trial to Evaluate Competncies through an Information Systems Education, IPSJ SIG Notes, 2008 (81), pp. 15-22 (2008).

26) Takahashi, K., Ichikawa, H., Kubota, S. and Abe, A., Development and Evaluation of Collaborative Tourist Information System based on Goal-centered Sequence Model, IPSJ SIG Notes, 2010-IS-111 (14), pp. 1-6 (2010).

27) Utsunomiya, H., Ohhashi, H., Takahashi, I. and Miyagawa, H., Evaluation of Information Systems, IPSJ SIG Notes, 93 (90), pp. 31-40.

ACM (evaluation of IS)

28) Paris, C.L. and Colineau, N.F., "A Cost-Benefit Evaluation Method for Web-Based Information Systems," *HT'09*, pp. 353-354 (2009).

29) Dünser, A., Grasset, R. and Billinghurst, M., "A Survey of Evaluation Techniques Used in Augmented Reality Studies," *SIGGRAPH Asia '08*, (2008).

30) Ortiz-Hernández, J., Nieto-Ariza, E.M., Rodríguez-Ortiz, G. and Montes-Rendon, A., "A theoretical evaluation for assessing the relevance of modeling techniques in business process modeling," *SOQUA'07*, pp. 102-107 (2007).

31) Santos-Neto, P., Resende, R.F. and Pádua, C., "An Evaluation of a Model-Based Testing Method for Information Systems," *SAC '08*, pp. 770-776 (2008).

32) Hazlewood, W.R., Lim, Y.-K., Connelly, K. and Makice, K., "Exploring Evaluation Methods for Ambient Information Systems," *CHI 2008*, pp. 2973-2978 (2008).

33) Anderson, J.W., "Information Systems Performance Measurement and Evaluation," *Proc. ACM Computer Science Conf.*, p. 184 (1985).

34) Barn, B.S, "On the Evaluation of Reference Models for Software Engineering Practice," *ISEC'09*, pp. 111-116 (2009).

35) Jin, Y., Tang, A., Han, J. and Liu, Y., "Performance Evaluation and Prediction for Legacy Information Systems," *ICSE'07* (2007).

36) Nomdoe, H. and Pather, S., "Providing a Holistic Perspective to Evaluating e-Service Quality in Web-Based Systems: a Multi-Stakeholder Approach," *SAICSIT 2007*, pp. 97-105 (2007).

37) Mathieson, K. and Ryan, T., "The Effect of Definitional variations on Users' Evaluations of Information Systems," *DATA BASE*, pp. 37-48 (1994).

38) Bryant, K.C., "The Evaluation of Courses in Information Systems," *ACE2004*, pp. 15-23 (2004).

Effects of Action Comment Pictographs for a Video Sharing Website

Jun Munemori*, Kentaro Kagawa**, Junko Itou*

*Faculty of Systems Engineering, Wakayama University, Japan
**Graduate School of Systems Engineering, Wakayama University, Japan
{munemori, s105068, itou}@sys.wakayama-u.ac.jp

Abstract – We have developed low-impact comment-posting functions for operation on an actual video-sharing site, using two types of pictographs (feeling comments and action comments). We offer active action expressions (action comments) like "feed" by pictograph. In addition, we offer a function to extend the effect of the resonance sense and to prevent repeated comment posting (post-lock). We aim at the improvement of viewer satisfaction. As a result of experiments, we understood that there were numerous comments that were intended to induce positive action / a positive response. The "clap" pictograph of contributed a lot. Action comments seemed to be more effective than conventional comment functions. The resonance sense works well, presumably because feedback from the resonance sense promotes the contribution of comments. Concerning post-lock, the length of time (3 seconds) was felt to be slightly long for animation viewers.

Keywords: video sharing, pictograph, action comment, resonance sense, post-lock

1 INTRODUCTION

Most video-sharing site users tend to watch videos for enjoyment [1]. We developed low-impact comment-posting functions for operation on an actual video-sharing site [2]. Concretely, feelings that can be expressed frankly through animation were offered as a comment function. Video viewers can choose and post feelings from among the preset comment group through one-click operation. This function adopted the pictograph media because it is easy to recognize pictographs that work well on a video screen, and can control improper expressions. In addition, we included a resonance sensory function [2] that acts interactively with comments posted by other users, and performed an application experiment for continuous comment-posting.

In this paper, we aimed at environmental improvement to promote contribution of comments for animation viewers using a video-sharing site. In the past study of the authors [2], users requested pictographs for other than emotional comments. We improved the pictograph comment that we developed for the class of users who had not thus far contributed comments and offer active action expressions like "feed" by pictograph. In addition, there is a function to extend the effect of the resonance sense and to prevent repeated comment-posting. The overall aim is the improvement of viewer satisfaction. This study will contribute to a study of the communication that assumed animation mediation

2 RELATED WORK

"Nico Nico animation" [3] is administered by "niwango" and is the greatest video-sharing site in Japan. Viewers can contribute their ideas on animation to the text area blanks. As for the contribution of comments, it is done very well on an animation screen. Viewers can thereby read comments while concentrating on the animation screen.

The contributed comments move from the right side to the left side over an animation screen. Comments are adjusted to specific animation scenes and displayed. Viewers share a sense of reality and are able to raise it.

However, in "Nico Nico animation," there are cases in which the visibility of animation is reduced by excessive text comments covering the entire screen, called "a barrage," or by using net slang such as "wwwww." Such cases are expressed with "rough weather", and are criticized.

Yamamoto developed iVas [4], a system that performs annotation work, which is necessary for semantic search and abstracts of animation content. When an animation reader performs annotation of animation in iVas, Yamamoto defines "impression annotation" as organization reducing the burden. Impression annotation is an interface to input the subjective impressions of the animation reader. Impression annotation consists of two steps using bar graphs which visualize an " impression button" inputting the subjective impression of the animation viewer such as "interesting" or "important" and the statistics on oneself and all readers. An animation reader can input the atmosphere of animation and one's subjective impressions by clicking buttons labeled "interesting", "sad", "strained" and so on. If we use the mechanism of iVas, it is difficult to simply adapt to the video-sharing site, as the same letters (interesting, sad, strained, and so on) can overflow on the screen. Therefore, pictographs are used rather than text (letters).

3 THE SYSTEM (FRIED ONION)

We have developed a video-sharing website system called Fried Onion.

3.1 Design policy

The function designed in Fried Onion aims at the development of a user-friendly comment contribution function. First, we expanded the kinds of pictograph comments that can be input easily. In addition, the limit function was created at the time of the comment contribution to prevent inappropriate comments. Finally, a resonance sense was added to realize exchanges of interactive reaction between viewers.

3.2 Development environment

Table 1 shows a list of software that composes this proposed video-sharing system. Figure 1 shows a screen of the system.

Table 1: Software constitution.

component	software	version
Web server	Apache	version 2.2
RDBMS	MySQL	5.1
Scripting language	PHP	version 5.3.0
View content	Flash	CS4
Video encoder	FFmpeg	rev. 25566

We use Flash to provide an interface to synchronize animation and comments. We use FFmpeg to extract a thumbnail image from a movie file in time to contribute animation. The site user can contribute animation / delete / view and hear or contribute a comment. The user performs data communication using an Internet browser, which is equipped with a FlashPlayer plugin.

3.3 Operation

Figure 1: A screen of the top page.

Fried Onion includes some functions other than a function to see and hear animation. The top page, from which all functions of Fried Onion can be accessed, is shown in Figure 1.

First, the user of the site accesses the top page (Figure 1) of Fried Onion using a Web browser. From the menu column (upper part) of Figure 1, users can use a variety of functions: new registration, login, logout, animation contribution and deletion of animation. In addition, there is a search function under the menu, which enables searching for animation that the user wants to see and hear using tags and categories.

The site user changes next to the animation seeing and hearing page by choosing the animation that he/she wants to see or hear from among the animation thumbnail group in Figure 1. An example of the seeing and hearing page can be seen in Figure 2.

Figure 2: An example of the seeing and hearing page.

When the site user changes to the animation seeing and hearing page, the data of the animation that the user chose are sent to a Web browser by a Web server, and streaming reproduction is carried out automatically. The user reads the upper explanation in the seeing and hearing page in Figure 2 and can search for related animation from a tag.

The animation interface provided by Adobe Flash Player is comprised of an animation screen, a replay and pause button, a search bar, replay time, a volume adjustment bar, and a pictograph comment function. The screen size of the animation seeing and hearing interface has the following dimensions: width 640px height 530px. The animation screen has an aspect ratio of 16:9, a width of 640px and a height of 360px.

3.4 Pictograph comment function

Pictograph comments consist of action comments (bottom) and feeling comments (the upper part of action comments) in Figure 2. A pictograph spreads when the cursor is passed over it. As for pictograph comments, they are displayed well on the animation screen and flow through the left of the screen. As preventive a measure against reduced visibility due to an increase in the number of comments, the viewing area was made relatively smaller. In addition, contributed pictograph comments are kept at any time by a Web server.

-Feeling comments

Feeling comments are based on an affective three-dimensional model based on the wheel of emotions of Plutchik [5]. By this system, we adopted the most important primary feelings such as fear, surprise, sadness, disgust, anger, anticipation, joy and trust. We made pictographs of eight kinds of feelings adopted in reference to Unicode 6.0. Figure 3 shows the eight kinds of feeling comments (fear, surprise, sadness, disgust, anger, anticipation, joy and trust, respectively).

Figure 3: The eight kinds of feeling comments.

We created a pictograph indicating "surprise" with an exclamation mark and "trust" with a heart, as seen in Figure 3. The background color of the pictographs in Figure 3 is based on the wheel of emotions.

-Action comments

It may be said that "promoting awareness" and "nodding yes" become the main clues to communication in conversational expressions in the network community [6]. According to the hearing investigation for viewers of "Nico Nico animation," there were comments other than emotional comments such as "greetings," "questions," "making friendly responses," "feed" and "sing" [7]. In addition, in the past study of the authors, users requested pictographs for "feed "and "scream" [2]. Therefore, in this study, we offer active pictograph comments to the animation viewers called action comment, which are separate from feelings comments. Pictographs of the action comments made in reference to Unicode are shown in Figure 4 (greet, question, attention, scream, feed, ok, and clap, respectively).

Figure 4: Pictographs of action comments.

- The post-lock function

We can contribute pictograph comments offered by this system through one-touch operation. Some people tend to express the strength of their feelings by repeated contributions of the same pictograph. Therefore we developed the " post-lock" which was a function to control the contribution of repeated comments in the system for a short duration. An example of the post-lock screen is shown in Figure 5.

When a viewer contributes a pictograph comment, the post-lock is automatically carried out. The post-lock consists of an illustration of a clock and a green icon in the top left-hand corner of each pictograph comment, as seen in

Figure 5. While post-lock is in effect, all pictograph comment input operation becomes invalid for three seconds.

Figure 5: Example of the post-lock screen.

-The resonance sensory function

Animation viewers often contribute very similar comments. In addition, in Fried Onion, the kinds of pictograph comments that we can be contributed when seeing or hearing animation is limited to 15 kinds. Therefore the probability that the same pictograph comment is chosen rises. By this system, we established the "resonance sense" function to emphasize the ties of impressions of animation that one contributed oneself and another person felt simultaneously. We use Figure 6 and Figure 7 to display the time sequence of the movement of the resonance sense.

Figure 6: Before the movement of the resonance sense.

Figure 7: After the movement of the resonance sense.

Figure 6 shows a situation in which a certain viewer A contributes an "applause" pictograph comment after

viewing animation (the "clap" icon in the menu in Figure 6 spreads a bit). The same "clap" pictograph comments are then contributed on the screen by other animation viewers. Therefore, the function of the resonance sense works. A scene showing the resonance sense can be seen in Figure 7. In the top right corner of Figure 7, the pictograph which viewer A contributed is displayed enlarged. The size that the icon spreads is in proportion to the number of pictograph comments of the same "clap" that was already contributed for the same scene. The size of a pictograph comment is usually 35 px, but it is enlarged in size by 25 % for each pictograph comment that is contributed for the same animation scene. On the other hand, pictographs that have already been commented on are displayed uniformly in 60 px by length and breadth. In addition, when completely new pictograph comments are posted, existing pictographs fade-out at the same timer. In addition, when completely new pictograph comments are posted, existing pictographs fade out for a fixed period of time. The conventional resonance sensory function [2] used to operate for two seconds before and after, which is four seconds in total when the same scene is commented on. However, the questionnaire results showed that the conventional resonance sensory function was insufficient. Therefore we redesigned the resonance sensory function so that it would operate more easily by extending it from four seconds to ten seconds in total.

4   EXPERIMENTS AND RESULTS

4.1 Experiment

Experiments were conducted to discover the effectiveness of each function shown in Chapter 3.

Subjects were 12 people in their 20s and 30s who were experienced in using the jointly-owned animation site. Each subject participated in experiment at a different time. We show information on the animation that subjects saw and heard in Table 2, the scene of the experiment in Figure 8 and the experimental animation in Figure 9.

At first the subject gets used to the operation of the pictograph comment function by viewing two animation samples. The subject sees and hears animation from the left to the right and also from the left to the right in the next row below. The number of animation samples is 18, shown in Table 2. Subjects were told before experiments to contribute their impressions of animation using the pictograph comment function. Subjects completed a questionnaire about the experiment and the pictographs after seeing and hearing the animation. Experiments lasted approximately one hour.

Figure 8: Scene of an experiment.

Figure 9: Top page of animation of the experimental experiment.

Table 2: Information on the animation that subjects saw and heard.

Title	Content	Editing	Genre	Number of seconds
A stop-motion animated cartoon (No.1)	A stopmotion animated cartoon showing a stuffed toy	Stop motion animation	Animation	17
I make it roughly (No.2)	Produce Jack-o'-Lantern by a pumpkin	There is text commentary, music	production	163
I will play with Lego (No.3)	Produce Lego block by slow-motion replay	There is a sound commentary, music	production	192
The subway musician (No.4)	A performance of an orchestra	No	Music	84
A collection of CGCM (No.5)	Drinking- water commercial using CG	Features CG editing, music	graphic	82
Bowling (No.6)	A collection of bowling techniques	There is music	Sport	80
Lekue cooking (No.7)	Commentary on cooking with limited cooking equipment	There is text commentary, music	Cooking	146
A pitagora switch (No.8)	Playing a physics operation game	There is English sound commentary, music	Game	441
Made a hamburger steak (No.9)	Hamburger steak with a "delicious stick"	There is text commentary, music	Cooking	164
An unexpected special ability (No.10)	Undoing the speed of the Rubik's Cube	No	Sport	72
Intellectual curiosity (No.11)	Putting Mentos in Coca-Cola	No	Action	53
A dancing parrot (No.12)	The parrot of a zoo making a mysterious move	No	Animal	143
"Mogi Kon" university festival (No.13)	An Ogiri panel	No	Comic	127
"Mogi Kon" university festival part3 (No.14)	An Ogiri panel	No	Comic	153
I tried Rio cart (No.15)	Playing a racing game	Sound commentary	Game	848
I did it with cool mint gum (No.16)	Dissolve chewing gum and mix it in a dish	There is text commentary, music	Cooking	240
An encounter with Poo-chan (No.17)	Give a cat bait	A letter commentary	Animal	132
Smashing a watermelon (No.18)	Splitting open a watermelon with a headbutt	No	Action	15

4.2 Experimental results

From the experimental data, the number of contributed pictograph comments was 1,239; 609 were feelings comments and 630 were action comments. We show the number of contributions according to the kind of pictograph comment in Figure 10, and the number of the pictograph comments by each subject in Figure 11. In addition, the number of pictograph comments contributed to each animation sample is shown in Figure 12. We assign numbers from user 1 to user 12 in the order that they participated in the experiment.

Numerical values mentioned in Table 7 by Table 3 are the means of each item of the evaluation of scholastic ability on a five–point questionnaire that subjects completed after the experiment (1:I strongly disagree, 2:I disagree, 3:I neither agree nor disagree, 4:I agree, 5:I strongly agree).
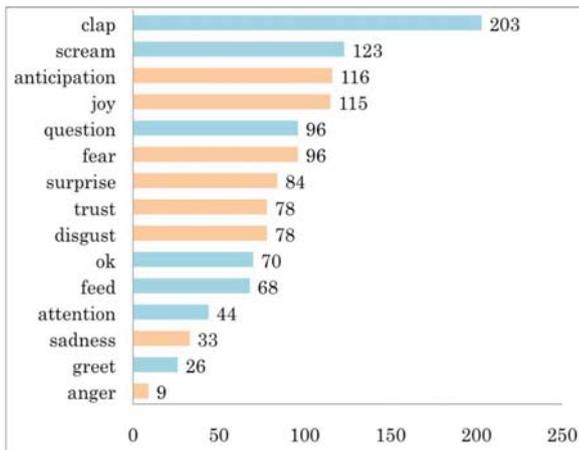
Figure 10: Number of pictograph comments contributed.
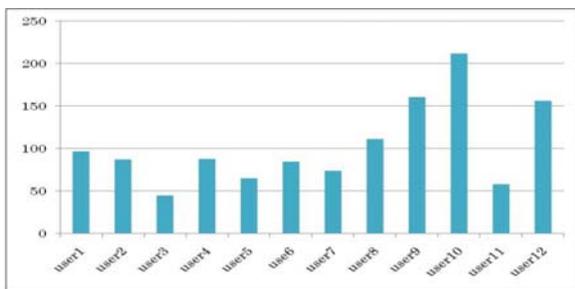
Figure 11: Number of contributed pictograph comments by each participant.
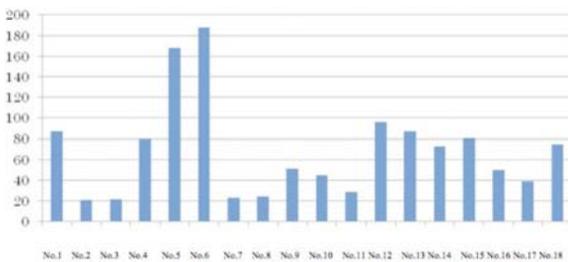
Figure 12: The number of pictograph comments contributed to each animation sample.

Table 3: Evaluation of pictographs.

Questionnaire item	Evaluation
The degree of availability of kinds of pictograph comments	3.6
Subjective satisfaction of the pictograph comment	4.3
Usability of the pictograph	4.5
Input time is shorter than normal text comment	4.8

Table 4: Evaluation of feeling comment and action comment.

Questionnaire item	Evaluation
Could you use the feelings comments well?	3.7
Did you think that the feelings comments were effective?	4.2
Could you use the action comments well?	4.2
Did you think there were sufficient action comments?	4.3

Table 5: Evaluation of the resonance sensory function.

Questionnaire item	Evaluation
I was able to confirm a response by the resonance sensory function and feedback	4.1
I was able to confirm the existence of other users	4

Table 6: Evaluation of the prevention of repeated comment function (post-lock).

Questionnaire item	Evaluation
I feel three seconds is a long time to wait to post a comment	3.5
I feel three seconds is a short time to wait to post a comment	1.8
I don't feel a need for a function to prevent repeated comments	2.3

Table 7: Evaluation of the communication with other viewers.

Questionnaire item	Evaluation
I contributed the same pictograph when I saw the pictograph of another person	4.4
I cannot understand which pictograph was which	2.5

4.3 Discussion

Pictograph comment

The mean number of comments per subject was 103, while 69 pictograph comments were given for each animation sample. Comments were posted at a rate of approximately one every 30 seconds, on average. It is 175 seconds for each animation reproduction time (mean value). Thus each subject contributed approximately 5-6 times for each animation sample. We can therefore confirm that it is a system in which it is relatively easy to contribute.

Subjective satisfaction with pictograph comments and evaluation of the usability were high in the questionnaire evaluation shown in Table 3. The item "Input time is shorter than normal text comment" was evaluated highly

(4.8/5.0) in this experiment. We found that it was easy to use the comments with two kinds of pictographs.

The evaluation of the item "The degree of availability of kinds of pictograph comments" in Table 3 was higher than the previous study (from 3.3 to 3.6)[2]. It is thought that the addition of the action comment accounted for this.

Resonance sensory function

The resonance sense is a function that allows sharing sympathy or an upsurge of emotion for animation with other viewers interactively at the time of posting the comment. A value of 4.1 was obtained for the item "I was able to confirm a response by the resonance sensory function and feedback" in Table 5. This exceeds 3.7, which was the value of the former study [2]. Only user 1, who was the first subject, gave 2 points on the 5-point scale (the resonance sense does not occur at first), but all the other users gave evaluations higher than 4. Because it lengthened the time that resonance occurs, the probability that a resonance sense was produced rose. Therefore evaluations also rose.

The item "I contributed the same pictograph when I saw the pictograph of another person" in Table 7 was high (4.4/5.0). This was confirmed from the animation which was shown in the scene of the experiment.

Figure 13 is a figure before a resonance sense occurred. Figure 14 is a figure after a resonance sense occurred.
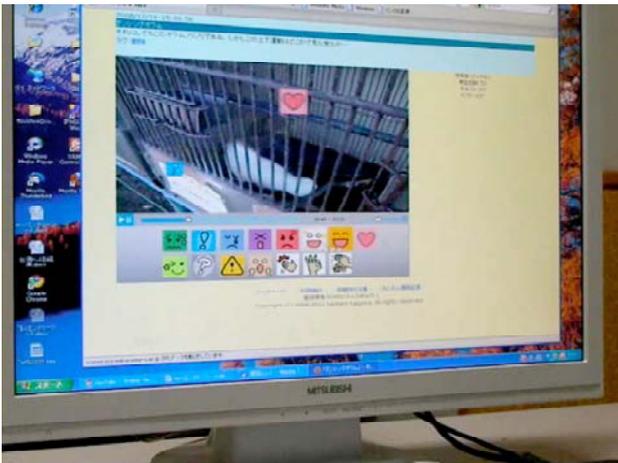
Figure 13: A figure before a resonance sense occurred.
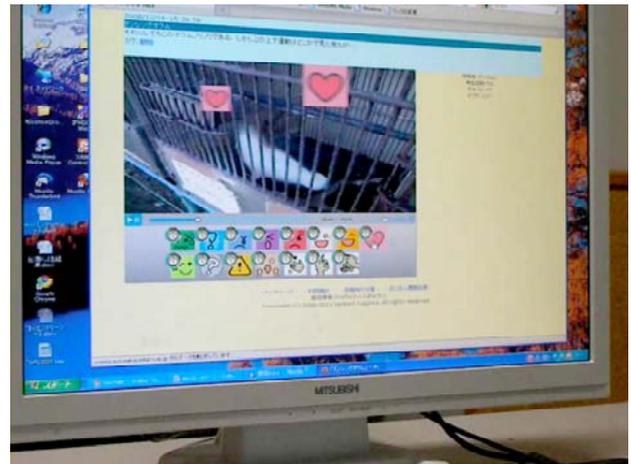
Figure 14: A figure after a resonance sense occurred.

The scene in which a viewer contributed a heart mark is shown in Figure 13, while Figure 14 shows a resonance sense taking place. The count of the resonance sensory function increases because comments are contributed not only for animation but also for other pictograph comments. Because comments are contributed for not only the content of the animation but also the contributed pictograph comments, the number of times that a resonance sense occurs increases. The resonance sense works well judging from Figure 11. For an experiment participant on the left whose user number is small, a tendency to use numerous comments that he/she contributed is seen in the user on the right (Figure 11). The mean number of comments of participants was 128.7 times from user 7 to user 12, whereas the mean number of comments of participants was 77.8 times from user 1 to user 6. In other words, resonance sense often works because the pictograph comment is accumulated as much as he/she sees and hears it towards the latter half of the experiment. Therefore, it is thought that it is one of the causes that the feedback by the resonance sense promotes the contribution of comments.

Action comment

Table 4 shows the division of pictograph comments into feelings comments and action comments. The number of times used and the effectiveness of action comments are superior to that of feeling comments. An action comment to show "clap" was contributed often (Figure 10). One scene of the "bowling" animation to which the "clap" comment was contributed particularly frequently is shown in Figure 15.

Figure 15: One scene of the animation "bowling".

The content of this animation is that a good player defeats a bowling pin by various throws; it appears that the viewer judged it to be good and contributed the "clap" comment seen in Figure 15. The "clap" pictograph is used in many ways. In particular, it was seen at the end of most posted animation scenes.

Prevention of repeated comment function (post-lock)

We discuss the prevention of repeated comment function "post-lock". The item "I feel three seconds is a long time to wait to post a comment" in Table 6 is evaluated relatively high (3.5/5.0) and the item "I feel three seconds is a short time to wait to post a comment" in Table 6 is very low (1.8/5.0). From the evaluation questionnaire, it may be said that it design was felt to be slightly long for animation viewers. By the description questionnaire result, there was an opinion that "three seconds is relatively long for seeing and hearing animation that was short". There was also the opinion that it was necessary as a means of preventing the animation from being damaged by pictograph comment.

5 CONCLUSION

In this paper, we aimed at environmental improvement to promote the contribution of comments for animation viewers using a video sharing site. We improved the pictograph comments that we developed for the class of users who had not thus far contributed comments. Action comments such as "feed" and "clap" were added to express the active action for animation scenes. In addition, we aimed at the improvement of satisfaction by improving the resonance sense and the function to prevent inappropriate comments by post-lock.

As a result of experiments, we understood that there were numerous comments that were intended to induce positive action / a positive response. The "clap" pictograph was often used. Action comments seemed to be more effective than the conventional comment function.

In addition, we proposed improvement of the resonance sensory function and a function to prevent repeated comments (post-lock). The resonance sense works well. Because it lengthened the time that resonance occurs, the probability that a resonance sense is produced rose. It is thought that this is one of the causes that the feedback by the resonance sense promotes the contribution of comments.

Concerning post-lock, the length of time (3 seconds) was felt to be slightly long for animation viewers. However, there was felt to be a need for post-lock

Improvements (for example, we add a sound to clap and a scream) will continue to be made in the future.

REFEENCES

[1] YouTube,Inc, http://jp.youtube.com/

[2] K. Kagawa, J. Itou and J. Munemori, "Effect of an Intuitional Pictogram Comment Function and an Emotional Sharing Function for a Video Sharing Web System", Journal of IPSJ (Information Processing Society Japan), Vol.51, No.3, pp.770-783 (2010)(in Japanese).

[3] Nico Nico Douga, http://www.nicovideo.jp/

[4] D.Yamamoto, K.Nagao, "Web-based Video Annotation and its Application", Transactions of the Japanese Society for Artificial Intelligence, Vol. 20, No.1, pp.67-75 (2005) (in Japanese).

[5] R. Plutchik, "Emotion: A psychoevolutionary synthesis", New York: Harper and Ro (1980).

[6] T. Nishida, "Towards Interactive Media for Community Knowledge Creation", IPSJ Magazine, Vol.41, No.5, pp.542-546 (2000)(in Japanese).

[7] T.Nakamura, H.Aoki, H.Miyashita, "NICOVIDEO Navigation System using Comic-like Visualization", Human Interface, Vol.10, No.4 pp.103-110 (2008)(in Japanese).

Life Log and its Application to Remote Consultation System

Yuya Uesugi[*], Jun Sawamoto[*], Norihisa Segawa[*],
Eiji Sugino[*], Takuto Goto[**], and Hiroshi Yajima[**]

[*]Software and Information Science Department
Iwate Prefectural University, Japan
[**]Department of Information Systems and Multimedia Design
Tokyo Denki University, Japan
sawamoto@iwate-pu.ac.jp

Abstract -Recently, opportunities that a consumer has to make choices are increasing. And, needs of the remote consultation is rising as a means of making choices efficiently. However, the burden of the expert (consultant) is large in the existing remote consultation, because there is a large difference between numbers of customers and experts. Therefore, it is difficult to carry out a consultation that considers the customer's personality and personal preferences and that matches well to the customer's needs. In this paper, we propose a technique to solve these problems by applying life log for basic data acquisition in remote consultation and present useful application examples.

Keywords: Remote consultation; Life log; Home electric appliance; Care and health consultation; Recommendation

1   INTRODUCTION

Recently, people's needs and values become more multifaceted by diversification of individual lifestyles. As a result, businesses offer various services tailored to customers' individual needs and their opportunities for choice are increasing accordingly. As the means to assist consumers' efficient selection of services, demands for remote consultation are increasing [1]-[3].

In existing remote consultation systems, it is, by its nature, difficult to propose a solution tailored to each individual customer. On the other hand, teleshopping system, for example, Amazon.com [4], can recommend products best fitted to each customer by learning the customer's latent preferences from the customer's checking and purchasing history of products and understanding the customer's characteristics of potential purchases. However, if you attempt a recommendation like this in current remote consultation systems, you ask and enquire preferences from people who are seeking advice or recommendation, and after understanding them from their responses, you recommend products accordingly, or you divide people into patterns by age, sex, yearly income, etc. and make recommendation accordingly. In any case, the procedure that you follow cannot be said better in both efficiency and accuracy.

When making recommendations for customers' individual needs, we need quantitative information that reflects the customers' personality and personal preferences. In the conventional remote consultation system, the main purpose is the problem solving, and it goes along mainly on dialogue between customer and consultant (expert). Therefore, the quantitative information that we can get is limited to data, such as age, annual income, and family structure, which does not reflect customers' preference and characteristics. Such fewness of the customers' personal data indices can be considered causing inefficient and less accurate recommendations of the existing remote consultation system.

In this paper, we intend to facilitate the consultation tailored to individuals, by collecting and analyzing life log from home electric appliances and estimating the latent characteristics of each individual. We will also further consider the application field where our consultation method can be effectively utilized.

This paper is organized as follows: Our preceding work of remote consultation system and related works concerning life log applications are reviewed in section 2. Then, we provide an overview of this research and describe the system configuration in section 3. The method of applying this research to several application areas is given in section 4. Finally, conclusions and the future prospects are described in section 5.

2   RELATED WORKS

2.1   Preceding study

In our preceding research [2[[5], we created remote consultation support system using outline generation agent in order to support interrupt and resume of the session in the process of consultation, and to facilitate better individual understanding for both customer and expert. In our research, we use outline to summarize the direction of the consultation, frequency of questions, transit time of consultation, and items that the customer wants to give higher priority by analyzing history of consultation process. Then the system visualizes the process of the consultation and makes the consultation task easier.

However there are some problems remaining with the system. For example, volume of information and efficiency is not sufficient enough because the system deals only consultation record as its analysis target and can get outline information only from customers' input, and also outline is used in limited situation and doesn't make sense in short consultation where interrupt and resume are not required.

2.2 Related study of life log

There are two patterns in study about life log, one study is about a method to get and reprocess life log and the other study is about a method for utilization of life log. Aizawa [6] [7] [8] realized an easy retrieval of life log video data using the query by combining life log video data with sensor data from GPS, gyro sensor, acceleration sensor, brain wave sensor and information database of weather and town etc. to assume context of life log and electronic document data. It is very useful to accumulate life log data combining with other relevant information and acquire life log data relevant to certain events easily. But, there are some problems when we apply life log video data to remote consultation. It can help to know the characteristics of individual customer, but it is a time consuming burden to expert because the data is made up of movie data.

Ito [9] achieved the improvement of the satisfaction and discovery characteristics of TV program recommendation by using the life log, such as viewing time of the TV programs, history of visited places, WWW browsing history. In life log made of "history" like this, there are chances that much valuable information is hiding because the customer does not need to keep a record consciously. If we could use such life log in remote consultation, we can change the current status of consultation where experts make a lot of questions to know the taste, the current state of the customer to the new status where experts acquire customer's data from life log and we can expect to realize more efficient remote consultation.

There are many specific researches to utilize life log to specify preferences of users [10]. But there is no research which applies life log data to remote consultation like our research. If we could use life log to remote consultation, we think we can scale up the range of consultation from cases that allow mistakes to some extent like recommendation of daily products to cases that allow no mistake like health care consultation.

2.3 Home network

Home network technology operates home electric appliances such as refrigerator, home video etc., through home LAN. There are two types of researches of home network. One is the research of the communication protocol between the consumer electronics, and the other is the research about applications that use consumer electronics through the network.

As the example of the former case, ECHONET (Energy Conservation & HOmecare NETwork) is proposed by ECHONET Consortium [11]. It utilizes sensors and controls and manages home electric appliances over network, aiming at facilitation of energy efficiency and home care services. In 2002, "ECHONET Specification version 2.11" was defined and in 2008, "Version 3.60" was issued and also registered as an international standard [12]. Such as automatic lighting on when entering a room is often adhere to this ECHONET standard.

As examples of latter case, Lin et al. [13] proposed UbiREMOTE, and Sekimoto et al. [14] presented BAMBEE.

UbiREMOTE displays the spatial layout of home electric appliances and home network on the remote control terminal to create a 3D virtual space, and the user selects the 3D graphics and operation menus to operate appliances by the centralized and intuitive manner. The user does not need to use many remote controls to operate various appliances. BAMBEE is a GUI system displayed on the touch screen. By simple operation, user can create and edit integrated services of home electric appliances. In this system, as well as for professional service providers, also for non-expert end users, it is possible to create applications easily. That can greatly expands the application range of integrated services of home appliances.

In this way, applications via the home network not only make the user's life more convenient but also promote efficient energy usage. However, these applications are mostly intended to send commands to the appliances via home network; we cannot find much study of handling life log and applying to various applications as we intend to do in our research.

3 PROPOSAL OF AN APPLICATION OF LIFE LOG TO REMOTE CONSULTATION

3.1 Outline of the research

Purpose of this research is to take advantage of personal information in daily life. We propose a system to get information of usage log from home electric appliances like TV, refrigerator, mobile phone and etc., then, to accumulate those logs through home network and to use it to remote consultation. We assume Android [17] based home electric appliances to access for getting life log. In addition, by using information obtained from these appliances, we discuss the application field of remote consultation support system to assist in the selection of products or service suited to personal preferences.

3.2 Life log

Typically, the word of life log means mainly video data like home video, and we use big machine like video camera for getting information. Consequently, it is hard to get life log easily from the point of mobility. However advancement of technology like mobile and wearable computer, and etc. enables us to get information with handy devices, e.g., a mobile phone with camera and GPS function. And services that gather and organize acquired information on the network as life log are provided from some providers already [18]. However we must make the log consciously like photo shoot in almost all of these services, and such data is fragmentary depending on the log-taker and log-taker may abandoned some data arbitrarily. Recently, it came to be able to acquire log information without user's consciousness by the appearance of the Android consumer electronic, and the method of using the life log extended greatly. Examples of life log that can be obtained from home electric appliances are shown in Table1.

From home electric appliances with communication function like mobile phone, life log concerning an individual activity history like the sojourn time in a stay place and a specific base etc. can be acquired. From the consumer electronic for the amusement, life log concerning the characteristic of contents that the person watches, the contents type of the favor, and media, etc. most often used can be acquired. Moreover, from the life consumer electronic, a potential characteristic what the person values in life for food and clothing etc. and the life time pattern can be presumed. If the mentioned information can be used for the remote consultation, the difficulty to grasp individual characteristics can be solved.

Table 1: Life log examples from househould electric Appliances

Type of electric appliance	Name	Life log Data
Communication	Mobile phone	Daily activity log
Amusement	TV	Watched TV programs, Watching hours
	Radio	Listened Radio programs Listening hours
	DVD	Watched DVD software Watching hours and recording information
	CD	Listened music Listening hours
Daily life	Refrigerator	Used hours
	Washing machine	Used hours, weight, wash type
	Microwave	Used hours, cooking menus
Others	Car navigation	Mobile history

3.3 Acquisition of life log from consumer electronics

As an example of life log acquisition from consumer electronics, NTT DoCoMo has provided smart tap [15] and Zojirushi Co. and NTT have introduced watching electric pot (i-pot) that sends a signal to a regional treatment center every time the pour button is pressed [16] etc. The former is plugged into an electric outlet in the house and used, by connecting home appliances to it, to measure the amount of electricity used. The latter is an air pot with wireless LAN to send a signal to the network when the pour button is pressed and the signal confirms to care givers that the elderly persons under their care are safe. Most of these existing products are achieved by giving some external devices. When information is acquired like this, information is not retrieved correctly in some cases. For example, if the device is got off by external shocks, until it is noticed and fixed, information cannot be obtained. Also, the system is weak for arbitrary falsification of information by the customer.

In our proposed system, we process by software as much as possible without using such external devices. Android consumer electronics process interrupt instructions such as pressing a button by installed Android OS. This means that

user applications can detect the interrupt instructions to get all the history of buttons used. By using external devices only where not using button operations, the aforementioned information leakage, tampering can be minimized, and it is possible to obtain information of life log close to individuals' real life.

3.4 Outline of the proposed system

We show framework of the proposed system in Figure 1. Home electric appliances send life log to cloud service regularly. Cloud service deposits life log data and arranges, processes according to the usage. When the expert requires, processed life log are once sent to the customer, and after customer's approval for the sake of privacy, the life log is forwarded to the expert. By using life log data in this framework, we can utilize remote consultation system applying life log data regardless of the place you stay.
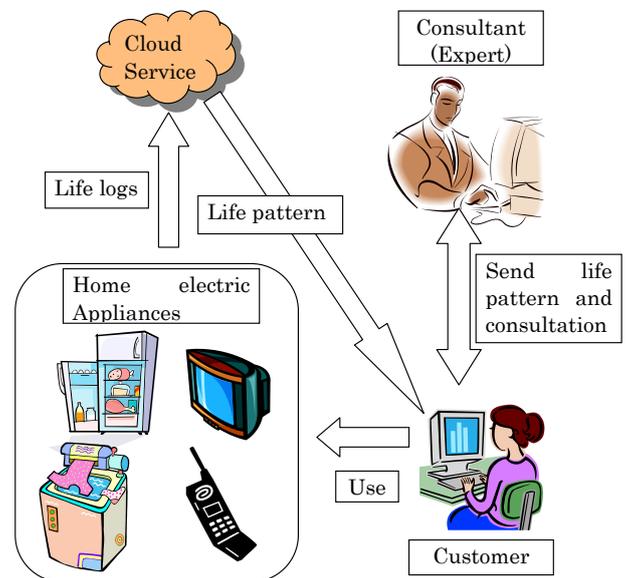
Figure 1: Framework of the proposed system

3.5 Process of life log

In this study, we get life log data from many tools. As a result, the volume of data that we treat in the system becomes very large and there exists a lot of data which identifies each individual. Therefore, when we use the data in remote consultation system, we have to process data and make it easier to apply to the application field (see Figure 2), and protect the privacy of the customer. The method of data process changes according to the target area of remote consultation. For example, when we apply this system to the area of elderly people's care, we can propose a care plan that match to the customer's life pattern, such as ordinary daily life pattern or specific occasion's life pattern, etc., from life log of mobile phone's mobile history, time record that the customer used home electric appliances first in the morning, etc.

As an example in the area of care support, the simplest way to create and use a life pattern is shown in Figure 2. Information acquired is as follows.

(1) Movement history from a mobile phone with GPS
(2) First and last time to use consumer electronics in a day
(3) Duration of usage of appliances related to cooking
(4) Duration of usage of appliances related to entertainment
(5) Duration of usage of bathroom equipments

From (1), the person's movement pattern and habitual outing can be derived. Moreover, a cycle to go out to see a doctor or the time to stay in health-related facilities can be acquired. Based on this information, the future planning for various activities including medical examination, etc. can be efficiently achieved. From (2), person's activity hours can be got. And personalized service can be provided by giving care support according to this time schedule. From (3), food and their time trends can be obtained. By analyzing the equipment used by the customer, you can guess whether the customer is having a healthy diet or the food is cooked properly. For example, the customer doesn't use any electric appliances other than refrigerators, or only the warm menu of the microwave oven is used, you can guess the problem with the dietary habits. As for microwave, because various recipes and variety of cooking function are given, you can get even more information. Based on the information, food policy can be constructed and it becomes possible to determine whether a diet-related services as nursing care should be recommended or not. From (4), the leisure time can be estimated. And health or care visits can be arranged smoothly. From (5), the time to bathe and the person is taking bath properly or not may be determined. Based on the information, whether the bathing care should be introduced, and when should be decided. From various kind of information, daily life patterns can be read, and patterns that differ significantly by the week or by the day of the week can be identified. Then by integrating those patterns and estimating the daily life patterns, various services can be created effectively for the customer.

4 APPLICATION FIELDS

In this section, we take up some of the application fields such as care and health consultation, travel planning consultation and consider about applicability and the convenience for the user in each field.

4.1 Care and health consultation

In Japan, needs of consultation related to physical condition of elderly people are rising from the effect of rapid aging society. For example, consultations about preparation of the care plan which includes what kinds of care support the customer needs, plans for care taker's visit schedule, etc (see Figure 3). The problem of the consultation is intractableness of getting necessary customer's information because customer is often old and sometimes information is taken from customer's family. They forget many things that are necessary for the consultation, and they don't report all the matter of necessity because they haven't enough knowledge of care. In some cases some people put on an impressive show. For example, they say "It is possible" regardless they can really do it or not. As a result, there are some cases that expert can't specify cause of bad health and the customer can't undergo appropriate treatment.

If we apply life log to care and health consultation, expert can get all the necessary information without customer's burden to record. If we are working on the care plan through consultation, the system can also check the degree of attainment of planned schedule using history of usage of home electric appliances. In this way, when we add life log in care and health consultation, we can realize high quality service and more adapted service to the customer.

Time	Action
5:40 am	Get up
7:00 am	Breakfirst
10:00 am	Outing
1:00 pm	Lunch
6:00 pm	Dinner
7:00 pm	Bath
10:00 pm	Go to bed

Life-logs   Life Pattern
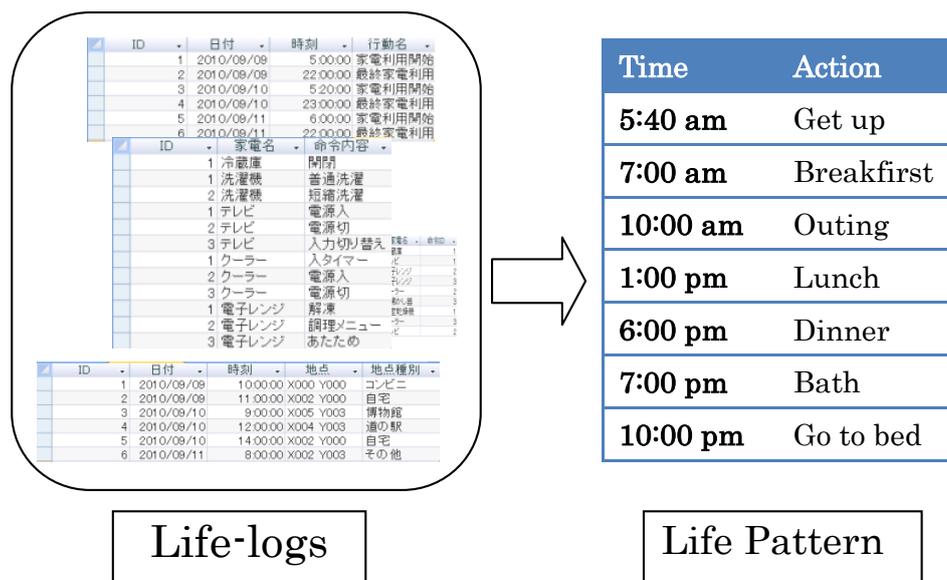
Figure 2: Life pattern deduced from life log data

Figure 3: A sample of care plan for elderly people

4.2 Travel planning consultation

We can travel more easily by the spread of service for the trip arrangement using internet. However, researches of travel planning consultation are increasing [19]-[23] because needs of travel consultation for customers who have no specific destination or have specific destination already decided, but have no detailed plan are increasing more widely ever. But in these preceding researches, user's preferences are extracted from some direct enquiries to the customer or certain patterns are identified from customer's age, sex, annual income, etc. and recommendations are given based on the patterns (see Figure 4). Preceding systems require much effort to match to individual customer's preferences and still hard to achieve detailed consideration about personal preferences such as sojourn time at certain visiting place.

By applying life log data to the consultation, the customer can deliver his preference without complicated input. Therefore, system can make the travel plan that match to customer's preferences without forming patterns of customers. And if the system gets stay time at per kind of facilities from GPS data of mobile phone, system can match the sojourn time of the destination in facilities to the individual's preference, and a comfortable travel plan without hurry and spare time can be set up.

Figure 4: Data enquiry window for a travel consultation system

23

5   CONCLUSION

In this paper, we presented how to get life log using the home electric appliances and examples of application field such as care health consultation, travel planning consultation. As a result, we could understand that it is very useful and it has a multiplicity of uses in consultation that closely attached to individual life, and in decision that based on individual preference. In the future, we will follow trends of home network and the development of Android-based household appliances further on to pursue feasibility of the proposed concept and system. We will continue to consider effective method of life log data processing and check the further applicability of life log data from the home electric appliances to remote consultation by constructing the actual application systems.

REFERENCES

[1] Y. Terahama, H. Kojima, H. Mizuno, H. Kinukawa, "Consultation System with Contact Person through On-demand Calling of a Proper Expert for a Problem," Journal of IEEJ C Vol. 126, No. 4, pp.457-464, 2006.

[2] H. Yajima, J. Sawamoto, K. Matsuda, "Remote Consultation System Using Hierarchically Structured Agents," HCI International 2009, July, 2009.

[3] Customer Service Optimization (CSO), http://www.artificial-solutions.com/

[4] Amazon.com, http://www.amazon.com/

[5] Y. Uesugi, T. Goto, N. Sakakura, J. Sawamoto, H. Yajima, "Development of Remote Consultation System Using Outline Generation Agent," Proceedings of the 2010 IEICE General Conference, B-19-12, 2010.

[6] K.Aizawa, S.Kawasaki, D.Tancharoen, T.Yamasaki, "Efficient retirieval of life log based on context and content," ACM Multimedia Workshop on Continuous Archival of Personal Experience 2004, pp.22-31

[7] K.Aizawa, "Capture and Processing of Life Log: Wearable, Ubiquitous, Automobile," The 19th Annual Conference of the Japanese Society for Artificial Intelligence, 2005, 3A3-01.

[8] K.Aizawa, "Processing of Experience : Capturing and Retrieval of Life Log," IEICE Technical Report PRMU , Vol.103 , No.738, pp.1-9, 2004.

[9] T.Ito, Y.Nakamura, H.Tezuka, S.Muto, M.Abe, "A Study on Personalized TV-Program Recommendations Based on Life Log," ITE Technical Reports vol.33, No.37, pp.81-86, 2009.

[10]H. Tezuka, K. Ito, T. Murayama, S. Seko, M. Nishino, S. Muto, and M. Abe, "Restaurant Recommendation Service Using Lifelogs," NTT Technical Review, Vol. 9, No. 1, 2011. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201101fa6.html

[11]ECHONET CONSORTIUM , http://echonet.gr.jp(Jan 13,2010).

[12]S.Yamada, "The Outline of the ECHONET Consortium(Services Achieved by the Force of Home Netowork<Special Issue> The Home Network)," The journal of the Institute of Image Information and Television Engineers Vol.59, No.5 , pp.705-709 (2005).

[13]Y.Lin,S.Yamamoto,M.Tamai,T.Kitani,N.Shibata,K.Yasumoto,I.Tomoya, "Remote Control Framework for Operating Networked Home Appliances via 3D Virtual Space," IPSJ SIG technical reports, vol.2008, no.18,pp.9-16(2008).

[14]J.Sekimoto,M.Nakamura,H.Igaki,K.Matsumoto, "Supporting End-Users for Creating Integrated Services in Home Network System," IEICE technical report. Information networks 107(525), 289-294, 2008-02-28.

[15]Eco-life in a Smart Apartment Room with an Integrated Smart Tap Network : http://i-energy.jp/data/Smart-Apartment2010-09.pdf

[16]Zojirushi i-pot http://www.mimamori.net/service/index.htm (in Japanese).

[17]Android Developers, http://developer.android.com/index.html, (July 14,2010)

[18]Evernote Corporation, http://www.evernote.com/about/intl/jp/, （July 14,2010）.

[19]Y. KURATA, "Toward an Interactive Tour Planning System," http://www/kurata/research/YKurata-GISA09.pdf,（2010/07/14）.

[20]Y. Kurata, "Challenges in User-Adaptive Tour Planning Systems." In Tomko, M., Richter, K. eds. AGILE Workshop on Adaptation in Spatial Communication, 19-26, 2009.

[21]Y. Kurata, K.Okunuki,Y.Sadahiro, "The development of preference-based tour planning system," Papers and proceedings of the Geographic Information Systems Association 9, 199-202, 2000.

[22]F. Ricci, B. Arslan, N. Mirzadeh, A. Venturini, "ITR: A Case-Based Travel Advisory System." In Craw, S., Preece, A. eds. ECCBR 2002, LNCS 2416, 613-627,2002.

[23]J. Lee, E. Kang, G. Park, "Design and Implementation Planning System of a Tour for Telematics Users." In Gervasi, O., Gavrilova, M. eds. ICCSA 2007, LNCS 4767, 179-189,2007.

Synthetic Research for the Information Security of the E-business & Recent Cloud Computing

Masaru Makino[*], Tadanori Mizuno**

* Graduate School of Business Administration and Computing Science (Doctoral Program),
Aichi Institute of Technology
**Faculty of Information Science, Aichi Institute of Technology

Abstract –I had put forward "E-business Security" about 10 years ago in Philadelphia IRMA (Information Resource Mgt Association). Nowadays we hear "Cloud Computing" in the world everywhere. Then I added "Cloud Computing Security". However, Cloud Computing is insufficient and incomplete in these days everywhere. In the big theme, I must research to the Cloud Computing. Now in this journal, the author has accepted only recent limited Cloud Computing. I suppose that I can put forward the publication for near future research.

1. GLOBAL SECURITY RULE FOR INFORMATION SYSTEMS

1.1 Information Security

The meaning of "Information Security" is that we keep the information assets from crises which threaten the information assets. The information worthy assets for enterprise that is which we can utilize information certainly and safely.

1.2 6 Elements of Security [10]~[13]

We have 6 security elements those are3+3=6 elements. The former 3 elements are fundamental and the latter 3 are added elements.

(A)Fundamental 3 elements
(1) Confidentiality
　　Only the person in charge can certainly access the required information.
(2) Integrity
　　The system secures security and completion.
(3) Availability
　　The users who are permitted can access information and relational assets at the necessary time.

(B) Added 3 elements
(4) Accountability
　　Who are users? What is service? What is responsibility? Service systems should keep user ID, operations history, logging history, evidence by signature etc.
(5) Authenticity
　　Contents are correct. Contents are not false. Users and systems behavior are clear.
(6) Reliability
　　Systems and process behave with no contradiction.

These 6 Security elements are international rules those are made by next standardization group.

　BS (British Standard Organization)
　ISO/IEC (International Standard Organization)
　　　/ (International Electrotechnical Commission)
　JIS (Japan Industrial Standard)
　ISMS (Information Security Management System)

1.3 Sample Technological Words of Security

The security technologies in 6 elements are separated by next technology group for instance.

Cipher or Code, Network or Web, Business Application Software, Middle Application Software, Basic Software, Cloud Computing, Virtual Engineering, General Hardware

These groups have many security technological words. The author says that this journal paper has Words books in backward reference [13].As for security systems, we have much security software such as Firewall - soft, PKI-soft, VPN-soft, Busting Virus-soft, Digital Money soft, and so on.

We utilize this software as for encrypting business-document, digital certificates, detecting privacy, tracing hacking-route, seeking business-scam, access control and restriction, securing electronic money, Packet filtering, discovering spoof etc. And then in these days, we must treat new security systems in the Cloud Computing.

2. AN INSTANCE OF E-BUSINESS (B to B)

2.1 Security Systems in Payment Activity

As shown in Fig.1, E-Business server- computer supervises business transaction-data with seller and buyer. Real seller and buyer in this system generally do not meet face to face mutually. The company or enterprise is often virtual systems. Therefore security information for seller and buyer is exceedingly important in these systems.

As shown in Fig.1, E-Business server-computer supervises business transaction-data with seller and buyer. Real seller and buyer in this system generally do not meet face to face mutually.
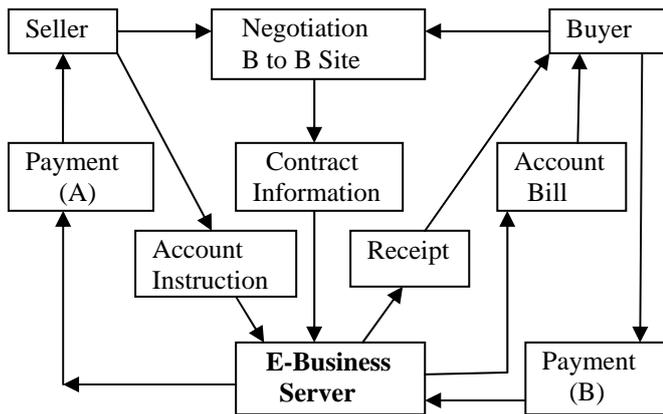
Fig. 1: E-Business (B to B).

Notes: E-Business Server has condition with credit, insurance, etc.

The company or enterprise is often virtual system; therefore security information for seller and buyer is often exceedingly important in these systems. Especially payment (B) and receipt are most important in Fig 1. We must be exceedingly careful in selecting any sort of electric and digital moneys those are E-credit, E-check, E-transfer, and any E-money(IC-card, Cyber cash, etc.) This electronic and digital money are recently called high-tech money which involves cellular phone & electronic wallet and other mobile & ubiquitous unit. We utilize certification system for the purpose of certifying seller and buyer whether he is real being or false being. False being is called as spoof. It is not easy to perceive the truth in business transaction. However we must approach nearly risky situation for the purpose of getting profit. Generally we often hear that business management is challenge to the risk. Risk on the business activity often occurs as the criminal act of money trouble.

Therefore the author suggests that we should prepare the payment systems as much as various sorts of cases. Then we must select the payment system according to the significance of business relationship. Moreover the author says that E-Business has relationship with E-Bank E-Finance, E-Commerce, E-Auction, E-Government, E-Insurance, E-Community, and so on.

2.2 E-Business (B to B) using UML (Unified Modeling Language) [24] ~ [27]

The author says that we can use UML for development of OOS software (object oriented system software). OOS has functions of "Object /Method /Message" & "Class/ Instance/ Inheritance". We can use these functions for making E-Business software System. These functions are attached to each box in Fig. 2 E-Business (B to B). The functions of OOS are composed those are System/ Super class/ Class/ Object. Systems are all boxes in Fig.2. Super class has plural boxes of class. Class has objects. And class has Class Name/ Attribute/ Operation/ Method. Object has Object Name / Attribute/ Operation/ Method.
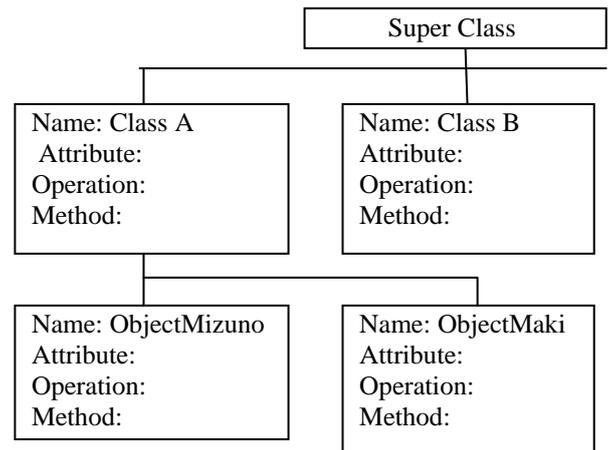
Fig.2: E-Business (B to B) using UML.

Notes: We can describe characteristics of those are Name/Attribute/Operation/Method in each box.

3. TRADEOFF -VALUATION APPROACH IN SYSTEMS DESIGN OF E-BUSINESS [1] ~ [3],[9]

3.1 Conceptual thinking of Tradeoff- Valuation Approach (TVA)

The author guesses that you know relationship between prices and employment in the economical society. If prices would be going up, then employment opportunities have more and more expanded. This relationship is approved in opposite situation. That is to say, if prices would be falling then employment opportunities have less and less become narrow, and as a logical consequence unemployment increases more and more. This relationship is defined as that A and B are in the tradeoff-relationship with each other.

In systems design, we expect much profit that is brought by the result of object-activity which is business transaction or systems model. However profit doesn't come true without security systems.

In this paper, the author suggests the next hypothetical formula in systems design.

Profitability = Feasibility – Security

Feasibility is expecting result of object-activity in systems design. Object-activity is brought from each system/ subsystem in which is exclusive of security systems. That is to say, feasibility is active technology on the other hand security is passive technology.

The author says that that would be able to see tradeoff-relationship between active technology and passive technology. So if we might reject passive technology then we could not realize expecting better result of active technology.

Table 1: Trade off Valuation Approach (TVA).

Sys Security	Sys/Sub A	Sys/Sub B	Sys/Sub C	etc	Valuation
C/ I /A	a m	b n	c o		
A/A/R					
Others					
P					

Notes 1: C/I/A is regulated as Security Policy.
   C (Confidentiality), I(Integrity),A(Availability)
Notes 2: A/A/R is regulated as Security Design.
   A (Accountability), A(Authenticity), R(Reliability)
Notes 3:Estimated numerical value (earnings or costs) is written in each box.
a,b,c,~ = Earning (in Feasibility), m,n,o,~ = Loss(in Security)
  Each row has similar meaning in each box.
Notes 4: P (Profitability)
Notes 5: Tradeoff-Valuation Approach (TVA) / Masaru Makino, All right reserved.

3.2 Procedure of Valuation Approach (TVA)

For the purpose of Table 1by value of sys/sub and cost of development for security system the author asserts that we must have procedure as follows.
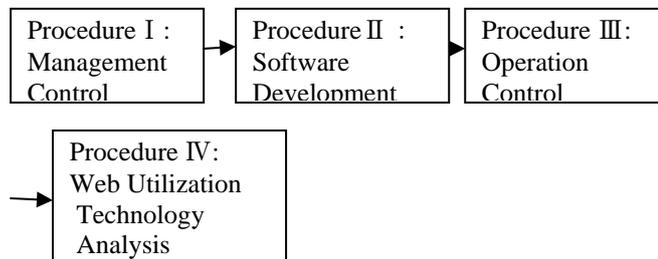
Fig. 3: Procedure of TVA.

Procedure I :This is the procedure concerned with management control. This procedure has sub-procedures, those are sales profit, sales amount, shared cost, cost of advertisement, and so on.
Procedure II : This is the procedure of software development. This procedure has sub-procedure those are the cost of web application software, security concerned software, and so on.
Procedure III : This is the procedure of system operation control. This procedure has sub-procedures those are outsourcing cost, organization measures cost, security measures cost, and so on.
Procedure IV: This is the procedure of web concerned technology analysis. This procedure has sub-procedures those are portal site analysis, top-page hit number, search

engine analysis access log analysis (key-word and access person analysis), web-site analysis (link situation, site design etc.).
The author asserts that we must use those procedure (I ～IV) for feasibility estimation of system/subsystems, and for security estimation of C/A/I or A/A/R in the Table 1. And moreover we must estimate other items for profit and loss situation.
We should always be aware of balancing for management and technologies those are great importance of trade-off valuation approach (TVA).

4. THE ENGINEERING TENDENCY OF RECENT CLOUD COMPUTING [15]~~[23]

4.1 Introduction on the Cloud Computing

The definition of "Cloud Computing" is that The NIST Definition is most popular and famous. NIST is National Institute of Standards and Technology in USA. The NIST Definition of Cloud Computing (Draft) is in that NIST Special Publication 800-145 (Draft).
The item of this Draft is COMPUTER SECURITY by Peter Mell and Timothy Grance, January 2011. In the item "1.Introduction "of the draft,"1.2 Purpose and Scope" says that NIST intends this informal definition to enhance and inform the public debate on cloud computing. Cloud computing is still an evolving paradigm. Its definition, use cases, underlying technologies, issues, risks, and benefits will be refined and better understood with a spirited debate by the public and private sectors. This definition, its attributes, characteristics, and underlying rationale will evolve over time.

4.2 5 Characteristics of Cloud Computing

Cloud Computing has many characteristics. In Cloud Computing, we can use big scale computer and get various services with small financing.
However, there are many curious cases in using cloud computing. The author has investigated through Amazon red book of cloud guides. On that case, the author has understood that next 5 characteristics of cloud computing.

 On-demand self-service
Access through the Broad Network
The pool of IT Resources
Scale in and Scale out
Automatic control and suitable utilization of IT Resources
Then the author says that most important events in the cloud systems are balance of power1 (demand and supply).
Next important events are balance of power2 (engineering and economy). And there are third important events etc.

4.3 Instance of Cloud Services / Cloud Computing

Cloud Engineering, Cloud Services, Cloud Computing are divided into many groups, those are next examples.

Table 2: Cloud Service/Cloud Computing.

	System	Software	Comment
SaaS	Force.com Amazon/EC2 Windows Azure (Others)	Gmail Java Script Lang Ajax	SaaS Service depend SYS
PaaS	Google AppEngine Force.com Amazon EC2/S3 Windows Azure (Others)	Script Lang Ajax Ruby Python Java	PaaS Service depend SYS
IaaS	Amazon EC2/S3 Google App Engine IBM Brute Fujitsu	Script Lang Ajax Java PHP	IaaS Service depend SYS

Notes 0: SaaS/PaaS/IaaS are Service Models in the NIST Definition of Cloud Computing
Notes1: SaaS (Software as a Service)
Notes2: Paas (Platform as a Service)
Notes3: IaaS (Infrastructure as a Service)

4.4 More Something around the Cloud Systems

We can see Table 2 SaaS/PaaS/IaaS around the Cloud Systems. The author had used next figure (3Step chart).

The author has understood that Table 2 is just Table 3. SaaS/ PaaS/ IaaS are just corresponding Application/ Middle/ Base in Systems Design.

Table 3: Systems Design Layer.

Virtual Enterprise
I **Application Layer** Design Planning/Sales Sys/Inventory Sys/ Purchase Sys/Receipt & Payment Sys
II **Middle Layer** Web Service/ Database/Security Sys
III **Base Layer** Platform/Server Operating Sys/Others
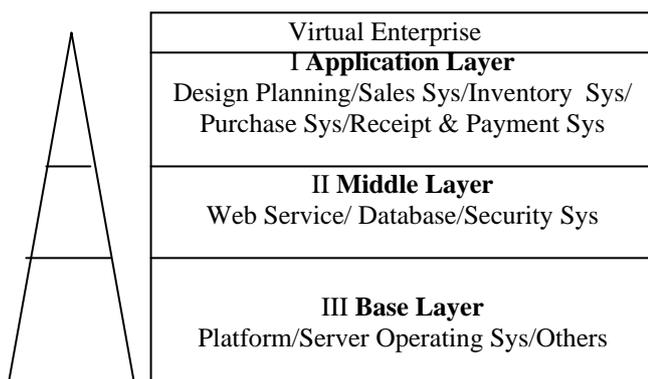
5. SECURITY SYSTEMS APPROACH

5.1 Systems Approach for the Security Systems Design

Systems approach is a direction of systems design for information systems development.

As shown in Table 3, information system is generally composed of three layers those are application/middle/base layers.

To give a few examples of each layer, application layer has EC (Electronic Commerce/E-Business),online trade of stocks by many sorts of terminals and internet banking middle layer has inter/intra/extra-network, database and cipher, base layer has OS and many sorts of servers those are firewall/proxy/www/E-mail etc. Systems approach for systems design has two directions those are top-down(I － II － III)、 and bottom-up　(III － II - 1).

5.2 Systems approach for security systems design

In case of large system, for example big business systems we must adopt top-down approach (I － II － III). However in case of small or middle systems, we may adopt bottom-up approach (III － II － I) or random approach.

In each case systems approaches for security systems design　depend on scale of system and other special reason.

6. TECHNOLOGIES FOR THE SECURITY SYSTEMS DESIGN

6.1 Security technology

The author presents security for information systems design in Table 4. The author has categorized items of security technology.

Table 4: Security Technology.

	Computer	Network
S Y S	SecurityPolicy/Securityagent/ Server/Cipher/ElectronicMoney RSA/Hacking/CERT-CC	Internet/Hacker/VPN/Proxy Cert:CC/DNS/IPV6/Virus TCP-IP/Mobile/Authent
S O F	AntiVirus/PenetrationAttack MaruWare/Security Agent DHCP Server/VPN/Logging	Firewall/IP-masqueader SSL/MIME/SMTP/BIND PackeiFiltering/Spam

6.1.1 Attack Operation

Attack operation in Table 4 is "Hacking/ Penetration/ Spam". In classification of virus, There are file infection, system domain infection, compound infection, Macro virus.

6.1.2 Defense Operation

Defense operations in Table 4 are "Cipher, CERT, Secure ID, Firewall, VPN, Proxy, IPv6, Anti-Virus, Authentication, Packet Filtering, and Logging". In these operation, we have technologies those are application-dependent or application-independent. For example CERT or

authentication is application-dependent. The author presents example of defense operation in 6.2.2.

6.2 Cipher Systems

6.2.1 Algorithm of Public key Encryption [8]

The computer makes cipher by algorithm of pseudo-random number. The formula is shown as follows.

In this formula X is original text. Y is cipher of X. r and n is are public keys.

$X^r = Y(\bmod n)$

$Y^s = X(\bmod n)$

Example of digits for each variable is following.

X=2, Y=326, r=13, s=61, n=437

(Notes 1) Explanation of these numerical values:
 13 power of 2 = 8192,
 8192/437= 18*437+326
 Residual (326) =Y, This residual Y is crypt value.

（Notes2）In these congruence expressions the author has used equal symbols. This equal symbol should be congruence symbol. However, the author cannot discover the congruence symbols in his desktop. He apologize you "I am sorry".

6.2.2 Example of Cipher System
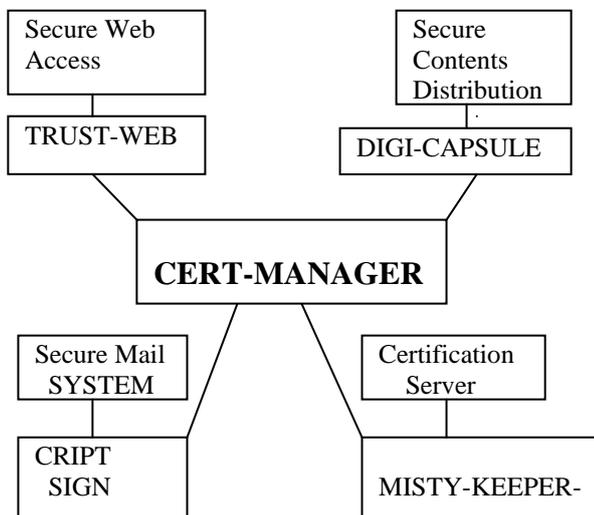
Fig. 3: MISTY Guard referenced by Ref [14]

MISTY is the encryption software that is made by Mitsubishi Electric Corporation. Mitsubishi encryption MISTY was progressing while 10 years or more over. Now MISTY is made from MISTY1 and MISTY2. MISTY is composed with encryption-design and block encryption (128bit key) and 64bit data block.

Encryption is important in secure and fast operation. You can see design, /Block cipher, /Differential crypt analysis,

/Linear crypt analysis /Provable security in the internet site (Block Encryption Algorithm MISTY).

7. THE CORRELATION BETWEEN CLOUD COMPUTING AND VIRTUALIZATION ENGINEERING [18],[19]

7.1 The purpose of Server Virtualization Systems

Virtualization system has notice that is one of important function for basic software engineering considerably ago.
And then, now we have high power computer system year and year. Virtualization system comes important year and year. Virtualization system is not captured by CPU, memory, network, storage etc., comes important higher & higher for basic engineering those are data center and cloud services.

Server virtualization system can move plural virtual machines those have each virtual operating systems.

Virtual machine can use many sorts of resources.

One virtual machine can divide and practice each submachine independently.

7.2 Two Types of Server Virtualization systems

We have two types of server virtualization systems those are host type and hypervisor type.

The host type system has server visualization software (VMware vSphere) and general OS (Windows & Linux etc.). The hypervisor type system has only server virtualization software and doesn't have general OS.

7.3 Instance of Server Virtualization

In the virtualization system we have ROI/TCO Calculation. These calculation expressions are ROI (Return on Investment) and TCO (Total Cost of Ownership).

By using ROI/TCO Calculation we can understand next various events of virtualization.
 (1) Cost cutting by server unification.
 (2) By utilizing virtualization-soft user can get effect in software development.
 (3) Business continuation and disaster recovery
 (4) Client management and security control
Notes: ROI/TCO Calculation expression is omitted here.

8. THE SECURITY OF RECENT CLOUD COMPUTING [4]~[7],[15]~[17],[20]~[23]

The author has selected 3 Groupings and Others for Cloud Security. 3 Groupings are NIST/CSA/IDC.

NIST is National Institute of Standards and Technology in USA Government (cf.4.1). CSA is Cloud Security Alliance. This is the ICT business group of famous corporations, e.g. DELL, Google, IBM, Intel, Microsoft, Novell, ORACLE, VeriSign, etc.IDC is International Data

Corporation Japan. IDC is in US, Asia, Pacific, and they publicize market forecast reports, e.g. Earthquake, Disaster, Smart city, Cloud Computing, Smart phone, Data center, Virtualization. Survey Results on Cloud Computing:

Anxiety to Security ①54.6%,
Anxiety to Response time and processing time:②26.3%,
Anxiety to Disaster : ③25.8%.

The author has selected another plural ICT magazine for Cloud Computing those are in backward REFERENCE (CLOUD SECURITY).

9. CONCLUSIONS

In the Opening on this JOURNAL, The author started GLOBAL SECURITY RULE that has 6 elements of securities. Next is E-Business and then the author presented his original research TVA (Tradeoff-Valuation Approach).

The author asserts that Cloud Computing is development-style from E-Business. However there is a dangerous bridge on the tradeoff scene too. Therefore, SECURITY & CRISIS are most important on the tradeoff scene too. We must how to research regarding such a difficult items, these are great theme. Table 1 in 3.1, there are Notes 1(Security Policy) and Notes 2 (Security Design). These 2 themes are biggest in coming issue for E-Business and Cloud Computing. These are new E-Business and new Cloud System in the near future. The author deals with these themes as for next research.

REFERENCES
GENERAL REFERENCE
[1]Tadanori Mizuno, Tetsuo Ideguchi, Takashi Okuda, Yoshimi Teshigawara, "Introduction of Computer Networks", Pearson Education Japan, pp.23~73(2001.1).
[2]Makoto Satoh, Masaru Makino, "The Keys for Success of SE Project", pp.62~94 (2000.3).
[3]Yoshiyuki Inoue, "Introduction to E-Business, Electronic Settlement System", Japan Zitsugyou Syuppan, pp.12~72 (2000.4).
REFERENCE (SCRIPT LANGUAGE)
[4]Hidenori Tsuji, Takashi Watanabe, Yukitoshi Suzuki, "Apache Web Server", ImpressJapan, pp.37~70(2006.7).
[5]Hershe, "PHP+MySQL", Hershe and Mainichi Communications,Inc, pp.111~146(2005.8)
[6]Kazuhiro Furuhata, "JavaScript", Mainichi Communications, Inc, pp.35~88(2003.11).
[7]Kazumi Takei, "Introduction to Web Application making by Ruby", Media-Tech, pp.140~187(2007.8).
REFERENCE (INFORMATION SECURITY)
[8]Masaru Makino,"Security Systems Design for MIS/OA related with Computer Network", 2000 MIS/OA International Conference, pp.205~208(2000.6).
[9]Masaru Makino, "Research on the Technology for Systems Design of E-Business", 2003 IRMA International Conference pp.1157-1159(2003.5).

[10][Hiroshi Kashiwado,"Base & System for Information Security ", Shuwa System, pp.10~22, pp.221~234(2011.2).
[11]ISO/IEC, JIPDEC,"Easy for Understanding of Information Security Management, ISO/IEC 27001Guide", pp.1~21(2011.4).
[12]ISMS, "Easy Understanding ISMS", NikkeiBP, pp.7~41(2002.8).
[13]NikkeiBP, "Technical Terminology for Security", NikkeiBP(2002.12).
[14] Mitsubishi Electric Corp.:Power MISTY(2000).
REFERENCE (CLOUD ENGINEERING)
[15]Changhwan Lee,"Amazon Cloud Technical Guide", ImpressJapan, pp.1~28, pp.165~195(2010.5)
[16]Tomoaki Abo, Takao Kobori, Yohsuke Kobayashi,"Development of Force.com Cloud Application", ImpressJapan, pp1~26(2010.4).
[17]NikkeiBP, "The Complete Cloud Computing", NikkeiBP, pp.1~97(2010.7).
REFERENCE (VIRTUALIZATION ENGINEERING)
[18]Ichiro Shibutani, Shinji Fukutome, ,Reiko Sawade, Masami Kudoh, "VMware vSphere4", ImpressJapan, pp.15~34(2010.7).
[19]ITpro/ NikkeiComputer/ NikkeiNetwork, "All Understanding Virtualization Daizen ", pp.1~28(2010.11).
REFERENCE (CLOUD SECURITY)
[20]NRI(Nomura Research Institute), Secure Technologies, "Information Security on Cloud Era", NikkeiBP, pp.1~43(2010.8).
[21]Nikkei NET WORK, "Introduction to The Newest Security Steps ", Nikkei NET WORK, pp.1~83(2010.8).
[22]Shintou Teramoto, Others Total 12, "Law Business on Cloud Era", Business Law Co..pp.1~50(2011.1).
[23]Peter Mell and Timochy Grace, "NIST Definition of Cloud Computing (Draft)", NIST(2011.1).
REFERENCE (UML(Unified Modeling Language))
[24]Sou Sagawa, Sakurai Tatsuya "Course of UML (Unified Business Modeling)" Soft Research Centre, pp.31~123(2008.6).
[25]Sinan Si Albir "UML Quick Reference",O'Reilly, pp.47~136(2000.2).
[26]Tohru Katou, Masahiro higuti, Shirou takada,"Object Oriented Java" Kindai kagakusya, pp.18~166(2010.8).
[27]Hitoshi sasaki,"Honkaku Gakusyhu Java", Gizyutu Hyouronsya, pp.129~202(2005.10).
REFERENCE (PROTOCOL)
[28]CLOUD SERVICE
http://www.google.co.jp/
Amazon/Salesforce.com/Hadoop/GoogleAppEngine//WindowsAzure/ROMA/MapReduce
[29]ICT Corp.
Microsoft/Google/IBM/Fujitsu/Vmware/NTTgroupe/ Nokia/Intel/Samsun/Apple/Panasonic/NEC/KDDI/ Yahoo/Softbank/Rakuten/Sony
[30]MobileCLOUD
Android/ iPhone/ Skype
[31]SECURITY Corp.
VeriSign/ itpro.nikkeibp.co.jp/security/ Trend Micro USB Security/MicrosoftSecurityEssentials

Session 1:
Home and Ubiquitous Networks
(Chair: Tomoo Inoue)

The development of database system for route recommendation

based on sensor data

Yoshihiro Ito[*], Yoshitaka Nakamura[**] , Yoh Shiraishi[**] and Osamu Takahashi[**]

[*] Graduate School of Systems Information Science, Future University Hakodate, Japan
[**] School of Systems Information Science, Future University Hakodate, Japan
{g2111003, y-nakamr, siraisi, osamu}@fun.ac.jp

Abstract - In recent years, we have become able collect granular sensor data through many sensors attached to a sensor network in the real world. This technology has been used for urban sensing. It is expected to achieve new personal navigation on the basis of sensor data. However, to recommend routes on the basis of sensor data, we need to handle road data and sensor data, but it is difficult for only the application side to handle these data. Therefore, a database system has been developed that can map sensor data to the road data by using spatial interpolation and provide path planning results to a navigation application. In addition, the effectiveness of the proposed system is discussed.

Keywords: sensor network, urban sensing, database system, spatial interpolation, route recommendation

1 INTRODUCTION

In recent years, miniaturization and upgrading of the sensor devices have been miniaturized and upgraded further. Because of this, we can gather environmental information in the real world easily. These data are measured at specified time intervals and stored in a database for applications. For experiment, there have been attempts that gather environmental information in urban areas. This is called urban sensing [1].

Urban sensing is expected to realize new personal navigation based on sensor data. For example, by sensing local crowded and high temperature places preliminarily, it is possible to direct people to a convenient route that avoids such places [2].

To recommend route on the basis of sensor data, we need to map sensor data to road data. At this time, sensors need to be searched for in the sensors in the vicinity of the road. However, if no sensor is in the neighborhood, it is difficult to find the route based on sensor data. In addition, if sensor data and road data are managed in different data format, it is difficult for the application side to inquire and process data.

For these reasons, in this study, we developed a database system that can map sensor data to the road data by using spatial interpolation and provide results of route recommendation to a navigation application. In addition, we discuss the effectiveness of our proposed system.

2 RELATED WORK

This section discusses research related to our study. Specifically, we discuss the research of a sensor database system that provides interpolated sensor data for application in section 2.1. In addition, we discuss research on urban sensing in section 2.2 and research on navigation and path planning in section 2.3.

2.1 Sensor database system

To use sensor data collected by a sensor network, they must be processed. There is the technology that can do this processing in the field of database. This technology considers the sensor network as a virtual database to which it sends the query. There is some research into this technology [3]. This technology needs to understand the inquiries between the application and sensor network. Specifically, a function is required that accesses sensor network and feeds back the answer to an inquiry to the application user. This system is called the sensor database system [4].

There has been research into a sensor database system to feed back interpolated sensor data. Iwai et al. [5] and Ishii et al. [6] suggested a system that generates mesh structured sensor data by using spatial interpolation and estimate sensor data in each areas. However this system does not estimate sensor data on roads necessary for route recommendation.

In addition, Arai et al. [7] suggested the database system for time series interpolation query. When sensor data which application requires does not exist, this system can estimate sensor data at arbitrary time by using cubic spline interpolation. However this system does not estimate sensor data anywhere else.

2.2 Urban sensing

We provide examples of research into urban sensing. Takagi et al. [8] and Tobe and Kurata [9] gathered environmental information in urban areas by using a sensor network. In these studies, they used a fine-grained sensor network within a 600m radius in front of Tatebayashi City Station in Gunma, Japan. This sensor network consists of leaf node and sink node. The leaf node gathers sensor data and transmits them to the sink node. The sink node transmits sensor data to database.

In addition, CitySence [10] is a sensor network that monitors meteorological phenomenon and aerial pollution in urban area. This system expands a sensor node's battery life by connecting the sensor node to a street lamp. By this approach, City Sense is used for long-range research. For example, it can monitor environmental full-time "Public health" means how healthy people are.

2.3 Path planning

2.3.1. Basics of path planning

In this section, we discuss basic elements of path planning. Solving the optimum route problem requires a graph structure like that in Figure 1.
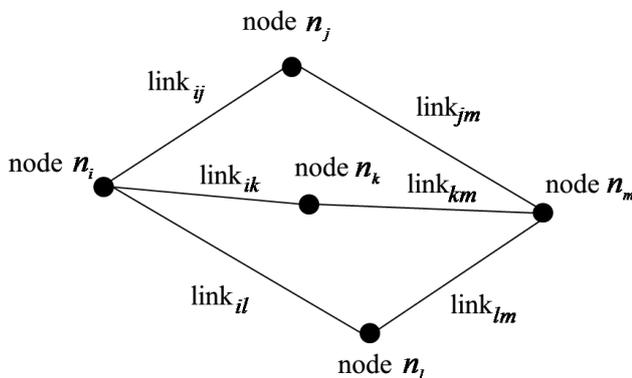
Figure 1: Graph structure

Each node is part of network that having information of node adjacent to them. In such structure, nodes are regarded as intersection, and the link as roads. This expresses the road network data. In addition, optimum routes can be searched for by giving the road distance to link as weight. Therefore, by changing one weight to another and paths can be planned on the basis of various parameters.

2.3.2. Related work on path planning

Architecture called pgRouting can search for optimal routes over the database [11]. This architecture can search for optimal routes by using a database table consisting of road data. This table needs a fixed-record format. It also needs "road id", "starting node", "end node", and "cost". However, pgRouting needs a table of road data. Therefore, when using sensor data as cost of the road data, a table must be created that records the road data for path planning, and updates it constantly during path planning. Applications have trouble executing these processes alone.

In addition, Endo and Tamura [12] suggested system using unusual WebAPI for navigation on the basis of sensor data. This system obtains sensor data by using this WebAPI, and calculates sensor data on roads, and searches route on the basis of these sensor data. This system is similar to our target system. The difference is that this system does not use special interpolation.

3 REQUIREMENTS OF PROPOSED SYSTEM

In this study, we aim to develop a system to recommend routes on the basis of sensor data. Therefore, we discuss its requirements below.

First, sensor data must be mapped to road data as cost. At this point, sensors need to be searched for in the neighborhood of the road. It makes no sense at all that there is sensor node on all roads. Therefore, an interpolate method is need to map sensor data to the road. In addition, our proposed system needs to calculate cost for route recommendation, to manage sensor data and road data at once and process inquiries from the application.

4 PROPOSED METHOD

4.1 Summary of proposed system

On the Basis of these considerations, we propose a database system to solve these problems. The database system needs the following functions.

A) Map sensor data to road data
B) Calculate sensor data-based cost of each road
C) Consolidate road data and sensor data
D) Process route recommendations
E) Process Inquiries

These functions are divided into two sections namely the Data Management Section which management various data and Database Management Section which recommends routes of the basis of sensor data (Figure 2).
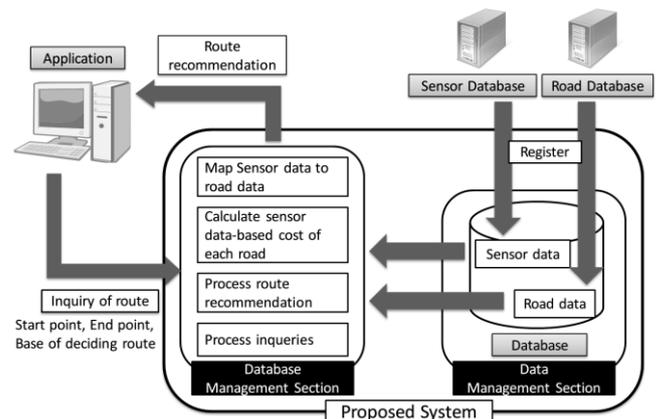
Figure 2: Summary of proposed system

The Data Management Section receives and manages sensor data and consolidates both road and sensor data. Road data are managed in road segment that separate each road from intersections. The Database Management Section receives sensor and road segment data and map sensor data to road segments. In addition, the Database Management Section handles and calculates the cost of road segment from mapped sensor data. Also, the Database Management Section processes inquiries and recommends route.

4.2 Management of road and sensor data

In this section, we discuss how to manage sensor and road data. Road data are managed as road segments that separate each road with intersections. Road segments are managed in the database as information of links described in section 2.3.1. Road segments are managed in the Road Segment Table which has the following data format (Table 1).

Table 1: Format of Road Segment Table

Attribute Name	Detail
Road ID	Identification numbers of road segment
Starting Node	Starting node numbers of road segment
End Node	End node numbers of road segment
Distance	Length of road segment (metric data)
Geometry	Location information of road segment (Line data)

The sensor data are managed in the Sensor Node Information Table and Sensor Data Table which has following data format (Table 2 and 3).

Table 2: Format of Sensor Node Information Table

Attribute Name	Detail
Sensor Node ID	Identification numbers of sensor node
Sensor Type	Types of sensor in sensor node
Geometry	Location information of sensor node (Point data)

Table 3: Format of Sensor Data Table

Attribute Name	Detail
Time	Identification numbers of sensor node
Data	The data which sensor node collected

Each of the Sensor Data Table named "Sensor Type_Sensor Node ID". In this way, we can access to Sensor Data Table by obtaining the Sensor Node ID and the Sensor Type from the Sensor Node Information Table.

The Data Management Section manages sensor data and road data as above, and the Data Management Section processes route recommendation by referring to all data in the Data Management Section.

4.3 Mapping sensor data to road data

Various factors that determine the environment of urban areas include buildings, parks and trees all of which affect the road environment. For example, when using temperature data, part of a road in the shadow of building is cooler than a part that is not. Therefore, to recommend the right route, sensor data need to be mapped on the road in a fine-grained manner. We explain how to do this.

First, we divide a road segment on the basis of its length. This is called a "divided road segment" (Figure 3). The number of divisions x is decided by the length of longest road segment in the road network (max_length) and defined width in meters is called "divide width" (d). Specifically, the number that is divided (i) is calculated as follows:

$$d \cdot (i-1) < x \le d \cdot i \ (d \cdot i \le max\_length, i \ge 1) \quad (1)$$

If the value of i is 1, the road segment stays undivided. In addition, the relationship between the road segment and divided road segment is as follows:

$$R_A = \{div_1, div_2, div_3, \ldots, div_i\} \quad (2)$$
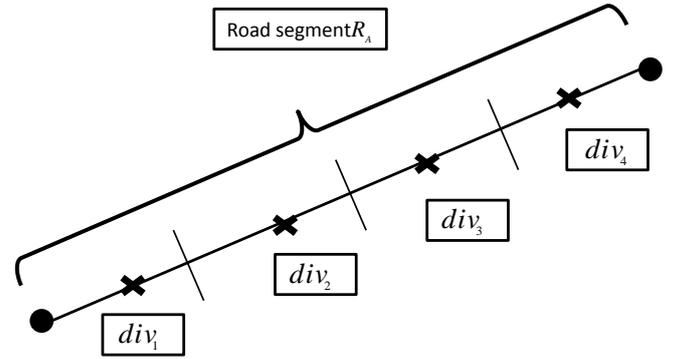
Figure 3: Divided road segment

By mapping sensor data to divided road segment generated by this method, we can know the environmental information on the road segment. However, depending on the situation of the sensor installation state, no sensor is considered to be near the road segment. In this case, possibly no sensor data are on the road segment, and this road segment can not be used for route recommendation. In our proposed method, we solve this problem by using special interpolation. Specifically, we estimate sensor data on a road segment by using Inverse Distance Weighting(IDW) [13]. IDW is the method most often used by GIS analysts. It estimates unknown measurements as weighted average over the known measurements at nearby points, giving the greatest weight to the nearest points. By using IDW, we can know the environmental information on divided road segment from a sensor near them.

When using IDW, nearby points must be found. The proposed method solves this problem by apply the Variable Radius Method [14]. The Variable Radius Method is the method used to draw data points by describing a circle with a radius of decided distance around an unknown point and locate data points in this area. When the required number of data point does not exist in the area of the circle, the radius is expanded to search for more data points again (Figure 4).

With the above method, we explain the method for sensor data estimation on a divided road segment ($div_1, div_2, \cdots div_i$). First, calculate the midpoint of divided road segments and set these midpoints as interpolate points. Next, define the search range in meters, and search for sensor node by using the Various Radius Method. Specifically, this method draws sensor node point by describing the circle with a radius of the search range introduced before around midpoint of a divided road segment. When no sensor node is in the area of circle, double the search range and search for sensor nodes again. It becomes possible to map sensor data on all divided

road segments by processing above method to the midpoint of divided a road segment.
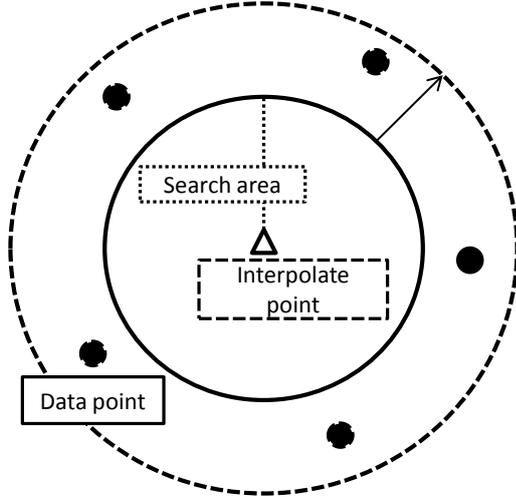
Figure 4: Variable Radius Method

4.4 Calculation of cost from sensor data

The method proposed in section 4.3, map sensor data on divided road segments. However, the cost needs to be calculated on each road segment for route recommendation. Therefore, we adapt the Dijkstra method which is the path planning algorithm because this algorithm excels in efficient and application, and define the cost-calculating formula for sensor data-based route recommendation. This formula calculates the cost of each road segment by using sensor value and the value which is decided on the basis of users feel comfortable (For example, using temperature, its value defined as 25 degrees Celsius). We explain this calculation formula below.

First, road segment R_A maps sensor data as shown in Figure 5.
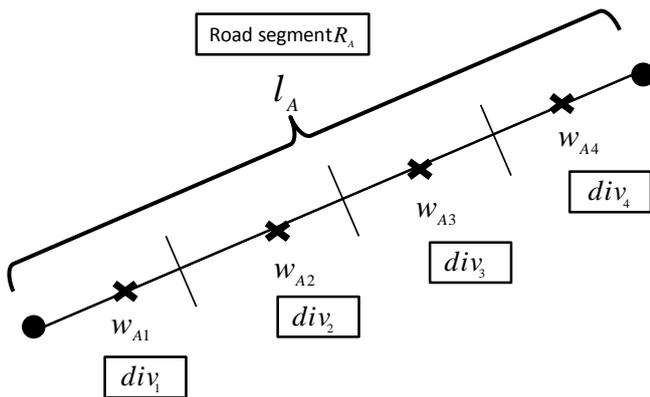
Figure 5: Parameter of road segment

w_{ik} is the estimated sensor data on divided road segment, and l_A is the length of the road segment R_A. In addition the division number of road segment is n, and the value which is decided on the basis of users feels comfortable as t_d. We find the sensor data based cost of each road segment by using these parameters. When sensor data and route length are considered, cost C_A is calculated by formula (3), and when only sensor data are considered, calculate cost C_A by formula (4).

$$C_A = \left\{ \sum_{k=1}^{n} |w_{Ak} - t_d| \right\} \cdot l_i \tag{3}$$

$$C_A = \left\{ \max(w_{Ak}) - t_d \right\}^2 \quad (1 \le k \le n) \tag{4}$$

Formula (3)'s cost of road segment increases as the difference between t_d and w_{ik} increases. Therefore, by using this formula, we can select a convenient route for the user. In addition, when the length of a road segment is multiplied by the difference between w_{ik} and t_d, cost C_A can consider the route length. On the other hand, Formula (4) attaches importance to sensor data. It does not consider the route length but calculates cost of road segment from sensor data only. Next, we describe formula (3) the "sensor and distance base method" and formula (4) the "sensor base method".

4.5 Flow of processing

Figure 6 shows the processing flow with the above method.
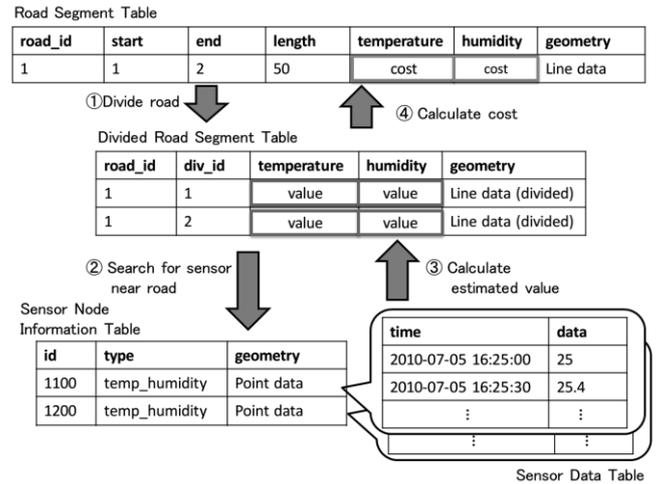
Figure 6: Flow of processing

First, the proposed method refers to the length of a road segment from the Road Segment Table and divides the road segment on the basis of the length. These results are stored in the Divided Road Segment Table (1). Next, to map interpolated sensor data onto a divided road segment, the method compares location information of the divided road segment and location information of sensor node, and pick up Sensor Node ID from Sensor Node Information Table by using the Various Radius Method (2). Therefore, by using Sensor node ID, the method searches for and obtains sensor data from Sensor Data Table, calculates interpolated sensor data on a divided road segment by using IDW and inserts these data into the Divided Road Segment Table (3). Finally, it calculates the cost for route recommendation by using the sensor and distance base method or sensor base method (4). After processing these methods, the proposed method recommends a route by referring to the cost in the Road

Segment Table and applying Dijkstra method. This data structure can control different types of sensor.

5 EXPERIMENT AND DISCUSS

5.1 Experiment environment

To verify the effectiveness of the proposed system, we conducted an experiment by using one computer (CPU: 3.2GHz，Memory: 4GB，HDD: 500GB OS: Windows 7 Professional). We developed a database system that had the function we proposed on this computer. Also, this database system is implemented by Java Servlet and uses a database-management system of PostgreSQL8.4, and it uses PostGIS1.5 an advanced function of PostgreSQL to manage geometry data.

5.2 Experiment for evaluation

In our experiment, we used sensor data gathered by the sensor network in Tatebayashi City, Japan. This sensor network gathers temperature and humidity data and a 600m radius in front of the station, as shown in Figure 7.

Figure 7: Distribution of sensor nodes

In our experiment, we used data gathered in summer, 2010. From these, we choose sensor data on clear days: August 23, 24 and 28, 2010. In addition, we used sensor data in August 21, 2010 because its temperature data on each point were high and low.

In our experiment, attempted to map and estimate sensor data. Specifically, we measured processing times elapsing from the beginning of the calculation to estimate the first to the last sensor data. In addition, we measured estimation accuracy in accordance with different times of day and evaluated the recommended routes. We discuss the experiment results below.

5.2.1. Route recommendation

We describe the result of route recommendation. In this experiment, we generated routes by using sensor and distance base method and sensor base method, and compared the generated and shortest routes. We used temperature data from 16:00 August 21, 2010. Figure 8 shows the results. Also, the value of temperature at which users feel comfortable (t_d) is defined as 25 degrees Celsius.

Figure 8: Generated route
(Left: Sensor and Distance Base Method
Right: Sensor Base Method)

To analyze these generated routes, we show a temperature gradient of road segment estimated by the proposed system in Figure 9.
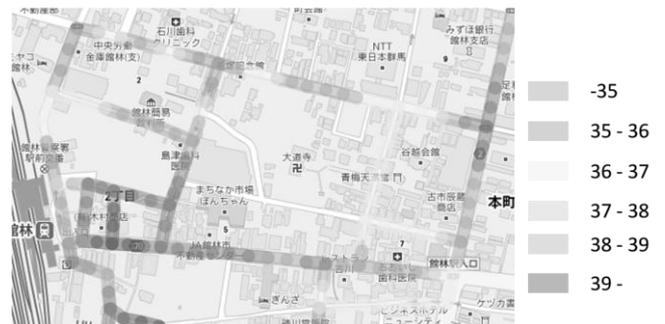
Figure 9: Temperature gradients of road segments

From this temperature gradient of a road segment, we define t_d as 25 degrees Celsius, and the proposed system was expected to recommend the coolest route. Figures 10, 11 and 12 show estimated temperature gradients of each route.
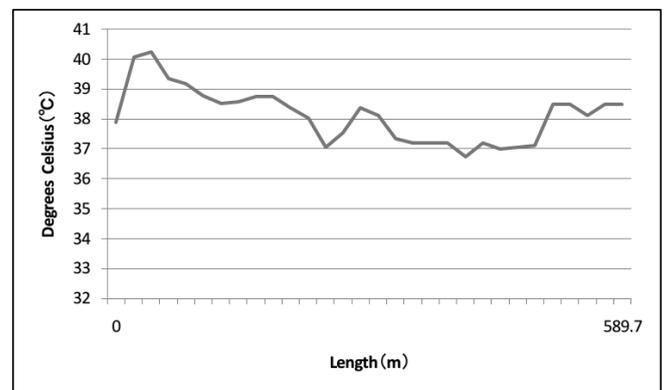
Figure 10: Estimated temperature gradient of shortest route
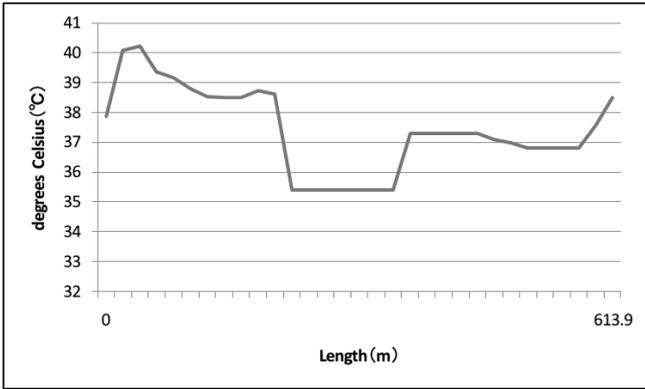
Figure 11: Estimated temperature gradient of sensor and distance base method route
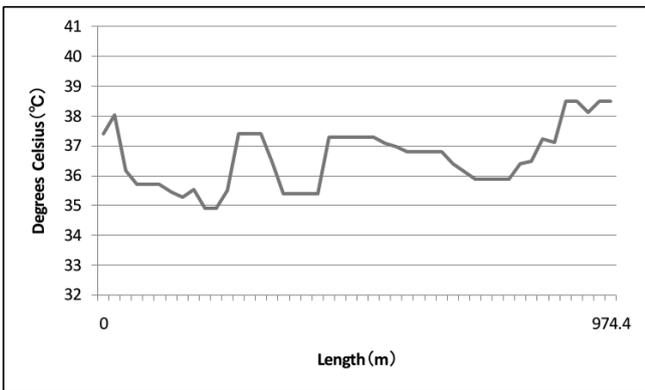
Figure 12: Estimated temperature gradient of sensor base method route

5.2.2. Processing time for estimating sensor data

We describe the processing time to calculate estimated value by using IDW. To use IDW, we define the value of the search range for the Variable Radius Method. We measure the processing time of value estimation by changing the search range. Specifically, we define the search ranges as 50, 100, 150, 200, 250, 300, 400, 450, 500, 1000 and 2000 meters, and measure processing time elapsing from the beginning of the calculation to estimate the first to the last sensor data of each range. These results are shown in Figure 13. Also, divide width of road segment (*d*) is defined as 25 meters.
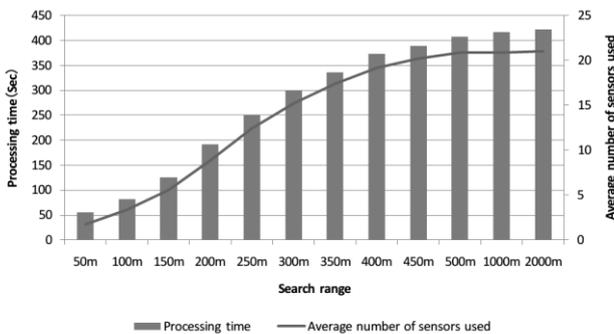
Figure 13: Processing time

According to these results, processing time increases as the search range expands, and the average number of sensor nodes used to calculate the estimated value at an interpolate point also increases. Processing time increases directly with the average number of sensor nodes used to calculate the estimated value, but when the search range is 500 meters, all sensor nodes are used to calculate all points, Neither processing time or the number sensor nodes significantly change after this.

5.2.3. Estimation accuracy

We describe results of an estimation accuracy experiment for the proposed system. This system has difficulty to analyzing estimation accuracy because it estimates sensor data on unknown points. For this reason, we set the place of a sensor node that has actual measured value in Tatebayashi City's sensor network as the interpolate point and estimated sensor value on that point by using proposed method. We tested this method on all points at which a sensor node existed, and calculated the difference between estimated and actual values. This difference was calculated by relative error and derived from the following calculation in percentage terms (E: estimated value, A: actual value):

$$\frac{|E - A|}{A} \cdot 100 \tag{5}$$

First, we measured estimation accuracy by changing the search range. Search ranges were the same as those given in 5.2.2', and we calculated relative error for all sensor node points by using formula (5) and calculated their mean. In this experiment, we used temperature data from 14:00 August 23, 24 and 28, 2010. Figure 14 shows the results.
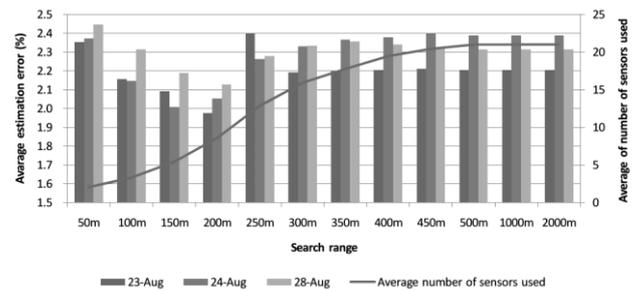
Figure 14: Change in estimated error by changing search range

As shown in Figure 14, as the search range expanded from 50 to 200 meters, estimation error decreased gradually. However, on the other hand, estimation error is increased from 250 to 500 meter. That difference may have been caused by using sensor data far from the interpolating points for estimation calculation. Next, we measured the estimation accuracy for temperature in accordance with different times of day. The times of day were 8:00, 14:00 and 20:00, and we calculated the average estimation error for each. Figure 15 shows the results.
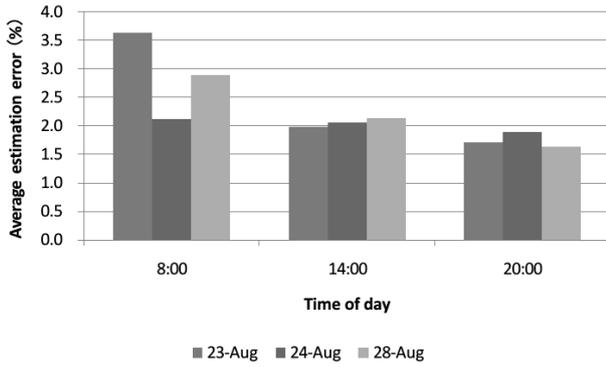
Figure 15: Estimation error by time of day

A 3% estimation error represented a degrees Celsius error. From this reason, Figure 15 shows the estimation error is small overall. However, at 8:00, estimation error was considerably larger on 14:00 and 20:00. To analyze this, we checked the estimated error of each estimated point at 8:00. Figures 16, 17 and 18 show the results.

Figure 16: Estimation error at 8:00 Aug-23
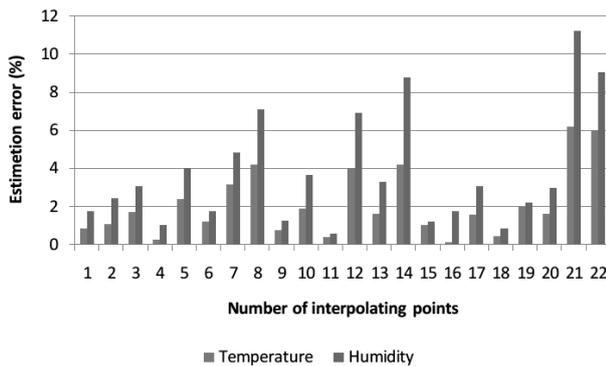
Figure 17: Estimation error at 8:00 Aug-24

Figure 18: Estimation error at 8:00 Aug-28

In this experiment, we calculated estimation errors for temperature and humidity. These results show the estimation errors for some interpolating points are large. In addition, this figure shows the degree of estimation error for temperature tends to be the same as the degree of estimation error for humidity. For example, we found that at places at which large estimation error was measured for temperature, a large estimation error was similarly measured for humidity. From these results, the estimation accuracy apparently depends on a sensor node's installation location and situation. Therefore, we analyzed the interpolating points at which estimation error for temperature was over 5% at 8:00 August 23. Figure 19 shows the results. Sensor nodes with a cross have no data or error data, so they were not used to calculate the estimation.

Figure 19: Interpolated point that measured large estimation error at 8:00 Aug-23

at 8:00 August 23, the installation location of sensor nodes the IDs of which ID is 18, 19, 22 and 12 had an estimation error over 5%. First, the standout feature of the interpolating point that had large errors was that they existed in the far end of the sensor network area. This applies to location of sensor node 18, 19 and 22. This reason is considered to be the lack of sensor nodes in far end of the area near these three. On the other hand, sensor node 12 exists near the center of the sensor network, but its estimation error was large. The reason is considered to be

that no sensor node for which data are available is near sensor node 1900, so it uses sensor node near another road to calculate estimations.

In another day's results, large error classifiable into these two patterns was measured at interpolating points. Therefore, the solution is needed to these problems.

5.2.4. Discussion of experiment results

In this section, we discuss results of the above experiment. In the evaluation of route recommendation, we confirmed that the proposed system recommend routes the way we assumed by using our proposed method and estimated sensor data.

However, from the evaluation of estimation accuracy, we found that some interpolating points have large estimation error. This error is possibly due to location and situation of interpolating points. This error affects the accuracy of recommended routes. To solve this problem, errors must be reduced at such points. One possible solution to this problem is participatory sensing [15]. By using participatory sensing, we can gather sensor data not only from a fixed sensor network but also from mobile devises. Therefore, estimation accuracy may be improved by gathering sensor data near point that have large estimation error by using participatory sensing.

In addition, from the evaluation of processing time, we determined that it takes time to estimate sensor data correctly. To solve this problem, processing must be speed up by improving database structure. Furthermore, the proposed system processes estimations and records this processing at constant intervals, and if an inquiry comes from the application, proposed system executes the route recommendation process only.

6 CONCLUSION

In this study, we developed a database system that can estimate sensor data on roads by using spatial interpolation and recommend routes on the basis of this sensor data. In addition, we evaluated processing time of the proposed system, estimation accuracy of sensor data and recommended routes generated by the proposed system.

In the future, to improve estimation accuracy of sensor data, we will consider a database system dealing with participatory sensing data. To achieve this system, we will consider a method to deal with mobile sensing data. Moreover, we will consider the way to use another sensor data.

REFERENCES

[1] Takahiro Ono, Hiroki Ishizuka, Kanoko Ito, Yasuyuki Ishida, Shohei Miyazaki, Oru Mihirogi, Yoshito Tobe, "UScan : Towards Fine-Grained Urban Sensing, Int'l Workshop on Real Field Identification, 2007.

[2] Keisuke Kanai , Takahiro Ono, Ishizuka Hiroki, Kanoko Ito, Yoshikazu Sawa, Oru Mihirogi, Yoshito Tobe, proposal of an Urban Information Sensing System, IPSJ SIG Technical Report UBI, No.18, pp.141-148, 2008. *(in Japanese)*

[3] Shigeru Ando, Yosuke Tamura, Yoshito Tobe, Masaki Minami, "Sensor Network Technology", Tokyo Denki University Press, 2005. *(in Japanese)*

[4] Yoh Shiraishi, "Database Technologies for Sensor Networks", Information Processing Society of Japan, Vol.47, No.4, pp.387-393, 2006. *(in Japanese)*

[5] Msayuki Iwai, ThePVILOJANAPONG Niwat, Hiroki Ishizuka, Nakamura Yoichi, Keisuke Kanai, Yoh Shiraishi, Yoshito Tobe, "TomuDB: A Database Management system for Distributed Urban Sensors", IEICE Technical Report USN, Vol.108, No.138, pp.13-18, 2008. *(in Japanese)*

[6] Nayuta Ishii, Yoh Shiraishi, Hiroki Ishizuka, Yoshiro Tobe, "An Integration Method of Heterogeneous and Multi-resolutional Sensor Data Utilizing Spacial Interpolation", Vol.49, No.SIG7（TOD37）, pp.44-58, 2008. *(in Japanese)*

[7] Kenji Arai, Yoh Shiraishi, Osamu Takahashi, "Design and development of a sensor database system for time series interpolation query", The 72nd National Convention of IPSJ (3), pp293-294, 2010. *(in Japanese)*

[8] Atsuhiro Takagi, keiji Sugo, Yasuyuki Ishida, Tatsuya Morita, Takeshi Iwamoto, Hideyuki Kurata, Yoshito Tobe, "A Practical Implementation of Micro-Climate Networked Sensing : Case at Tatebayashi City, Gumma", IEICE Technical Report USN, Vol.109, No.248, pp.13-18, 2009. *(in Japanese)*

[9] Yoshito Tobe, Hideyuki Kurata, "A Practical Implementation of Micro-Climate Networked Sensing - Case at Tatebayashi City", Information Processing Society of Japan, vol.51, No.6, pp692-699, 2010. *(in Japanese)*

[10] R. Murty, G. Mainland, I. Rose, A. R. Chowdhury, A. Gosain, J. Bers, and M. Welsh : CitySense: A Vision for an Urban-Scale Wireless Networking Testbed. In IEEE International Conference on Technologies for Homeland Security, 2008.

[11] pgRouting Project, "pgRouting",HTML Available at "http://www.pgrouting.postlbs.org".

[12] Tatsuro Endo, Yosuke Tamura, "Research of Sensor Networks by Using Web API- : Development of Navigation System by Using Sensor Data", The 70th National Convention of IPSJ (3), pp23-24, 2008. *(in Japanese)*

[13] P.A.Longley, M.F.Goodchild, D.J.Maguire, D.W.Rhind, "Geographic Information Science", John Wiley & Sons, Ltd, 2001.

[14] Zhang Changping, "Analysis of Spatial Data by Using Geographic Information System", KOKON, 2001. *(in Japanese)*

[15] J. Burke, D. Estrin, M. Hansen, A. Parker, N. Ramanathan, S. Reddy, M. B. Srivastava. "Participatory sensing", World Sensor Web Workshop, ACM Sensys 2006.

The Concept of an Agent-based Middleware for Smart Home Environments

Hideyuki Takahashi[†,‡], Taishi Ito[†,‡], Tetsuo Kinoshita[†,‡]

[†]Research Institute of Electrical Communication, Tohoku University, Japan
[‡]Graduate School of Information Sciences, Tohoku University, Japan
{hideyuki, kino}@riec.tohoku.ac.jp, itot@k.riec.tohoku.ac.jp

Abstract - In smart home environment, different kinds of system components including hardware elements, software components, network connections and sensors are required to cooperate with each other to reduce environmental burden by energy management and to support human's life allow for a comfortable lifestyle. This paper proposes a concept of an agent-based middleware for smart home environments that consists of various home electric appliances and various sensors related to smart grid. The agents acquire variety of information, data, etc. from the smart home environment and store/manage them in a methodical manner. Then agents configure and provide the energy management service, life-support service, multimedia service etc. based on the information and user requests. In this paper, we describe the concept, design and initial implementation based on our middleware. We show the initial implementation to confirm the feasibility and effectiveness of our middleware in the smart home environment.

Keywords: Multi-agent, Middleware, Smart home, Ubiquitous computing, Life-support system

1 INTRODUCTION

In recent years, researches on smart home [1] environments and service provisioning on the environment has been greatly accelerated. There are two mainstreams in researches on smart home; one is research on energy management system to reduce environmental burden, and the other is for life-support service construction scheme using various sensor information. In energy management system, researches proposal advanced methods for controlling energy and promoting energy saving to reduce environmental burden [2]–[4]. On the other hand, in researches on life-support service construction, superior frameworks and schemes are actively challenged for

Figure 1: An application of coordination of home electric appliances and various sensors in smart home environment.

dynamic cooperation among many kinds of system components, i.e., entities in smart home environments and ubiquitous computing environments, to provide user-oriented services [5]–[10].

The discussions on actual applications in smart home have been mainly focused on energy control of home electronics appliances and provision of life-support service, multimedia service, etc. In the future, these services control energy and appliances in user's house, and provide user depending on the situation based on various sensor information and user preference. In other words, we need to utilize various kinds of information to provide life-support service, multimedia service, etc. based on the infrastructure such as Home Energy Management System (HEMS).

We are promoting research and development on fundamental technologies aiming at smart home environment, including energy management service, life-support service, and multimedia service provisioning by cooperative behavior of home appliances and various sensors. We are targeting at the services such as home care support for elderly people who live alone including watching over, healthcare, safety confirmation, etc. from coordination of home appliances and sensor information as shown in Figure 1. We are also considering multimedia services such as video streaming or videoconferencing constructed from coordination of computers, home appliances, and smartphone considering Quality of Service (QoS), Quality of Experience (QoE), energy consumption, etc.

As for QoS and QoE, the system needs to satisfy user requirement in smart home environment. To address this, the system has to consider not only selection of the devices (appliances) from user location, but also resource situation of network, software, and hardware including electrical power. This is because resource availability tends to be poor and unstable depending on the device and status of use.

In this research, we are aiming to realize various service construction schemes in smart home in order to provide QoE-aware and energy-aware services against changes of resource status and user's situation. We have proposed an effective handling of multiple contexts including user context and resource contexts [11], [12]. To accomplish the objective, we apply agent-based middleware approach.

The concept of this approach is agentification of each entity in overall smart home environment. Agent has context management ability and cooperation ability for conflict resolution on multiple contexts. Agent also has maintenance mechanism for long-term context to accumulate and reuse history and experiences of past cooperation among agents. Agents would make the energy-aware and QoE-aware service provisioning
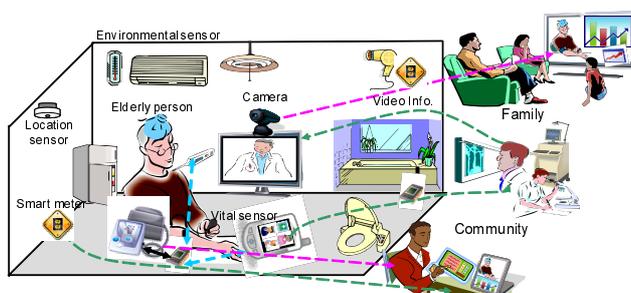
possible by the individual behavior and the cooperative behavior.

In this paper, we propose an agent-based middleware for smart home environment. Moreover, we describe design of the middleware focusing on the service construction scheme for QoE-aware and energy-aware service provisioning considering the multiple contexts. We evaluate our proposal from results of simulation experiments. We introduce the initial implementation of multimedia communication application based on our middleware to confirm the feasibility and effectiveness.

2 RELATED WORK AND PROBLEM

2.1 Related work

There are many middleware related to smart home and ubiquitous computing. The existing middlewares, frameworks, and service construction schemes are actively challenged for dynamic cooperation among many kinds of system components. CARMAN [13] considers Mobile News Service (MNS) for mobile users. It provides service based on user mobility, device's performance and user's preference. When the service is provided by a single mobile device, performance of the device is most important for the service. Therefore there are various kinds of works for providing high quality of web service within the limits of device's performance [14], [15]. These works are focusing on provisioning individual user-centric service using a single mobile device. To provide multimedia service by utilizing the available resources, we found it will be more efficient if multiple devices around the user can be used at the same time, instead of using only a single device.

In other similar works, service is provided by coordination with any devices around user [8], [10]. These frameworks construct the service based on service template which is requested by user. It means they search appropriate function to the requests and cooperation between various devices. Ja-Net [5] also aims to construct emergent service based on user preference.

These works' objective is same as our basic concept in terms of providing the service by coordination with heterogeneous entities. Moreover these works are providing superior mechanism of useful naming system, service description language, and service emergence.

However, we suppose existing service construction schemes are based only on user context and functional components, and they are concentrating on guaranty of coordination and operation or standardization of the specifications. It is important for smart home environment to satisfy a particular requirement and limitation including network resource, computer resource, and energy consumption. In case of rich service provisioning such as multimedia services in smart home environment, we suppose it is much more important to consider QoE. For instance, there are ever-changing situations like user mobility, device's performance around user, resource condition, and demand response based on energy management system. Therefore, there would be a possibility that devices that are physically very close to the user cannot provide

the service due to lack of the resources, even if the devices potentially have good performance. Moreover it's necessary to provide the service considering cooperation problem of unexpected devices, softwares, and network.

We concluded that it's required to achieve service construction scheme which considers not only user request but also other situation (context), in effective and integrated manners.

2.2 Problems

We need to address some technical problems to provide QoE-aware and energy-aware services on smart home environment that consists of computers, audio-visual home electric appliances, and sensors.

Resource context maintenance

We define "context" as situation of target entity at time t and temporal changes of the situation after/before time t. The situation is represented as internal representation model of the entity. Existing works have been mainly focusing on user context acquisition scheme such as users' locating information. However, in terms of resource of entity, it was treated as only a value of the target resource parameter at time t, not as "context". In smart home environment that consists of many kinds of entities in different level of functionality and performance, it is important to consider resource context efficiently as well for proper QoE control and energy management.

Multiple context coordination

In smart home environment on which heterogeneous entities coexist, QoE and energy consumption should be maintained in the range from entity level to overall system level. Therefore, we have to consider not only functional specification of the entity, but also multiple context coordination including resource context and user context.

Non-deterministic property of service construction

There are mutual dependencies and interoperability among entities that are not resolved deterministically from analysis of static specifications. Each entity is basically designed to work by itself, not designed to work with unknown entities cooperatively. Thus, services constructed from the entities would not work whether entities consistent with applicable specifications.

Effective acquisition of various and amount of information related to smart home

Many wired/wireless sensor devices detect environmental data and vital data in the smart home. For example, there are power consumption, location, vital sign, and brightness in real time. As for home care support service, the information has limitations for obtaining an accurate estimation because the information is obtained by the vital sign limited piece of information on certain individuals. It would be possible to perceive the health condition of elderly person with greater accuracy using physical location the person, environmental information such as ambient temperature, room brightness, energy consumption, and video information of the person, as well as the vital sign. However, it is difficult to acquire all the information because of the limitation of computational resources and network resources include wireless sensor's battery. Consequently, we need to consider the effective way of information acquisition.

Service provision based on various kinds of information

After acquisition of various kinds of information, effective information and service provisioning using the information would be a challenge. The data and information including energy consumption, vital sign, location information, environmental information, multimedia data, etc. contain significant diverse aspects in both quantitative and qualitative. Therefore, we need to construct the service provision mechanism include provision of required data and information, and control devices based on data and information.

3 CONCEPT OF AN AGENT-BASED MIDDLEWARE

3.1 An agent-based middleware

In this paper, we propose the following three approaches to solve technical problems.

Agentification of each entity

We define "Agentification" as a process making a target entity workable as an agent by adding knowledge processing mechanism. We also add context management ability, cooperation ability to resolve context conflict to the agents, and adaptive communication ability. Moreover, we embed long-term-context maintenance ability to the agents to accumulate cooperation history and experiences.

Multi-context-based Service Construction scheme

To realize QoE-aware and energy-aware service construction considering multiple context, we propose contract-based service construction scheme of agents. Agents make organization based on Contract Net Protocol (CNP) [16]. Moreover, we model heuristics and dependency information on cooperation history in past among agents as long-term context among agents. This kind of context is also managed by the agent. By using this context, agents can construct more advanced services employing lessons learned.

Control scheme of demand response based on user policy

To control energy consumption in smart home, we propose the control scheme of home appliances based on user policy and user's situation considering QoE, QoS, power consumption, and CO2 reduction.

The fundamental framework of our middleware is shown in Figure 2. Our middleware consists of four layers, i.e., Primitive Agent layer (PA), Agent Relationship layer (AR), Agent Organization layer (AO), and Smart Home Service layer (SHS). PA makes physical entities to agents. For instance, the agents have ability to manage context, control sink node of sensor network, selection of communication protocol based on kinds of data, etc. in this layer. In AR, inter-agent relationship based on long-term context among agents is created and maintained. In AO, agent organization is constructed based on the context in PA and AR, when user requirement or situation to specific service is issued. On the top layer SHS, actual service is provided to users.

3.2 Process of Service provision

Fundamental process to provide service consists of the following six steps.
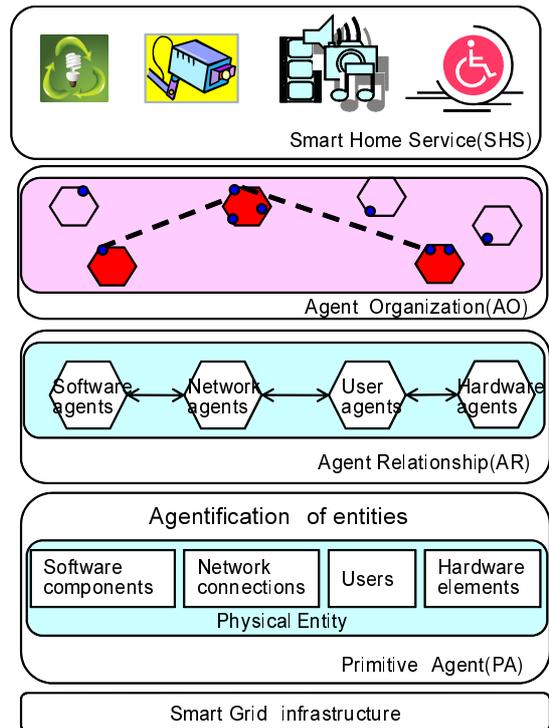
Figure 2: Agent-based middleware

(1) Agentification

Agent's designer adds each entity to domain oriented knowledge representation model that is suitable for classification of the entity.

(2) Updating of IAR

We define long-term context among agents as Inter-agent Relationship (IAR). Each agent has different IARs to all other agents with which it has cooperated in past time. Each agent updates their IAR after its service provisioning by itself or by cooperation with other agents.

(3) Self-directive user requirement acquisition

Agent considers and acquires user requirement autonomously. Agent analyzes the requirement from user's profile, locating information, and behavior in the smart home environment. Agent has to choose in the most suitable manner to get the requirement, because useful input devices may not be available in everywhere. This remains an open problem and is not treated in this paper.

(4) Service construction and (5) Service provisioning

To provide service, agents construct its organization using CNP. In this scheme, we apply hierarchical CNP, i.e., task announcement is propagated in order of hardware agents, software agents, and network agents. The agent organization is created based on context managed by each agent, and actual service is constructed and provided with combination of entities controlled by agents.

(6) User evaluation

The agent organization receives user's feedback concerning the quality of provided service when the service provisioning is finished. We introduce us-effectivity (E) based on [5]. In our middleware, when E value changes, each agent informs the update to all other agents that have relationship
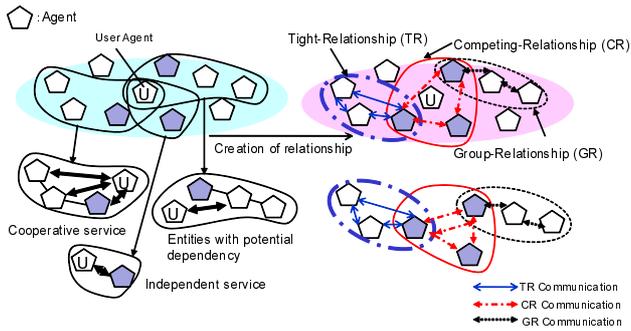
Figure 3: Creation of Inter-Agent Relationship

to it. Therefore, the result of evaluation is propagated to all related agents. It effects to service construction subsequently.

4 SERVICE CONSTRUCTION SCHEME

4.1 Basic Inter-Agent Relationship

Basic IAR consists of Tight-Relationship, Group-Relationship, and Competing-Relationship like shown in Figure 3.
(a)Tight-Relationship

Agents create Tight-Relationship (TR) when agents provide some services by constructing organization. It is possible for the agent to have past cases of successes and failures in cooperation by using TR.
(b)Group-Relationship

Group-Relationship (GR) is given to group of agents that have some potential dependencies. For example, there is GR among hardware entities such as sink node and source node of wireless sensor network, smart meter and appliance, desktop PC, speakers, and PC displays. It is possible that agent informs changes in their states frequently to the agents within the group by using GR.
(c)Competing-Relationship

Competing-Relationship (CR) is formed among agents that have same function. Why this relationship is introduced is that these agents would compete when task announcement of the function is issued. The competing agents routinely inform their status to each other, and they can make good organization effectively when CNP-based negotiation runs by using CR. Moreover CR has the effect of reduction of messages because agents which have CR send message considering other agents situation.

4.2 CNP-based service construction with IAR

Our scheme which is based on CNP builds agent organization using IAR. CNP is a mechanism to make contract relationship among agents by exchanging messages such as task announcement, bid, and award, shown in Figure 4. In this subsection, we briefly explain features of service construction scheme based on IAR.
(1)Case of TR

In Figure 4(1), we assume that agent A has a TR with both agent B and agent C whereas no IAR exists between B and C. TR between A and B indicates that trouble was occurred when

they had cooperation in the past, and TR between A and C indicates no trouble in the past. They refer to each IAR when B and C receive the task announcement from A. B does not send bid because TR against A is bad. That means the trouble in cooperation would occur this time too. On the other hand, C sends bid because C judges from TR that it would contribute to the task announced. It is possible to reduce trouble in cooperation by agent considering coordinated relationship in the past.
(2)Case of GR

We assume that agent A has no IAR with both agent B and agent C whereas relationship of type GR exists between B and C as shown in Figure 4(2). C recognizes that GR against B exists when C judges the task announcement from A. Then C sends bid if C judges that B can provide service by referring to state in IAR. On the other hand, C ignores the task announcement if B cannot provide service. It is possible to reduce the trouble in cooperation by agent considering dependency of the agents.
(3)Case of CR

In Figure 4(3), we assume that IAR of type CR exists between agent B and agent C whereas agent A has no IAR with both B and C. B and C receive the task announcement. Each agent checks IAR of type CR if it can process the task. When agent has CR, it refers to state of the CR. For example, B sends bid in case that B judged the value of us-effectivity on this task is higher than that of C. On the other hand, C ignores the task announcement in case that it judged us-effectivity of B is higher than C. In fact, it is possible to efficient construction of service by consideration of state of same function agent.

4.3 Policy-based home energy management

In general, demand response is a control scheme of power consumption of consumers on the electric power provider side. The scheme controls operational status of home electric appliances. We propose a policy-based home energy management depending on demand response as shown in Figure 5. For example, our proposal method controls the air conditioner about the preset temperatures during periods of peak demand for electricity, turns off air conditioners instead of switching on a fan cut electric power consumption, controls electric pot, microwave, illumination etc. to reduce wasted electricity depending on user's policy. On the other hand, when the power goes down due to a disaster such as earthquakes, our method controls the illumination, radio, television at a minimum power consumption considering the context (time, room brightness) and the rechargeable batteries. Additionally, when there is no possibility of demand response, our method provide multimedia service using a wide-screen TV and audio equipments based on user's policy. In fact, each agent calculates the power consumption and CO2, and agents control appliances based on IAR and user's policy.
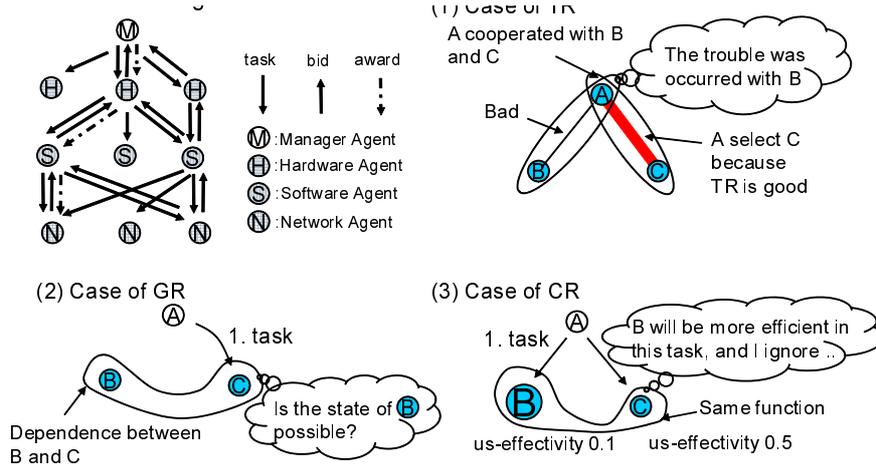
Figure 4: Characteristic of CNP by IAR
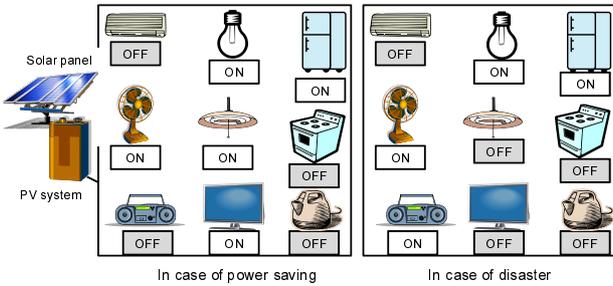
Figure 5: Overview of policy-based home energy management

5 SIMULATIONS AND EVALUATION

5.1 Implementation

To perform simulation, we implemented agents based on our middleware. We employed agent-based programming environment DASH [17]. We also performed simulation by IDEA [18]. IDEA is interactive design environment for Agent system. We used DASH because agent which is developed for simulation can easily be reused when we build the real-world system in future.

5.2 Evaluation method

In this simulation, QoE awareness of the system is measured and we investigate how much the QoE awareness is improved by introducing our middleware. To measure the QoE awareness, we apply User Request Achievement (URA) level. We can measure how much the user requirement is fulfilled by the system. Details of URA are described later.

Figure 6 shows the behavioral situation representation of the system. Here, three entities including a hardware entity, a software entity, and a network entity are making organization and providing service to a User. The user issues "User Request QoE" and the system provides service with "Provided QoE". Hardware Agent (HA) monitors CPU resource con-

text and Network Agent (NA) monitors bandwidth resource context. On the other hand, Software Agent (SA) has knowledge concerning mapping from resource availability onto actual user level QoE.

The QoE evaluation of service is based on URA. URA is calculated by comparison between User Request QoE RU and Provided QoE SV. In this simulation, we defined the range of URA is from -1 to 1. Here, ru_i is an element of RU and it represents User Request QoE on service element i. Also sv_i is an element of SV and it represents Provided QoE on service element i. The value of ru_i and sv_i is from 1 to 10. Here, URA on service element i, i.e. URA_i is represented as follows:

- $SV = \{sv_1, sv_2, ...\}$
- $RU = \{ru_1, ru_2, ...\}$
- $URA_i = (sv_i - ru_i)/10$

If URA_i is above zero, the user requirement is fulfilled. If it is below zero, the requirement is not satisfied.

In this evaluation, the number of service elements is assumed to be two ($i = 1, 2$) and URA indicates the total URA, that is, a mean value of URA_1 and URA_2 for simplification.

We performed simulation for 500 times. The agent constructs CR immediately after simulation beginning. When SA constructs service, it refers to NA's bid and IAR. If SA judges that other agent is suitable, it disregards the task even if the task is acceptable. And if agent receives bid by two or more agents that can fulfill user requirement, agent sends award to agent with the highest value of E of IAR after referring to the value of E. The simulation receives an assumed user evaluation each time of service construction. The user evaluation is reflected to E. User evaluation is assumed good if sv_i is within from 120% to 100% when ru_i is regarded as 100%. In this case value E is set to 1. It is assumed bad in case that sv_i exceeds 120% or is below 80%, and the value is set to -1. Otherwise it is regarded as usual, and the value is set to 0.

We compare three patterns of agent behaviors, i.e., our proposal (IAR-based approach), the case considering only user context (User-request approach), and the case considering only the maximum QoE value of agent for QoE without consideration of resource context (Maximum approach). Resources of HA and NA are assigned random values in every service construction. We also give tendencies of user request in four patterns, which is high quality (7∼10), middle quality (4∼7), low quality (1∼4), and random quality (1∼10).

5.3 Simulation results and evaluation

Figure 7∼Figure 10 shows the frequency distribution concerning URA. Figure 7 is a comparison for case that RU is always in high, Figure 8 is a comparison for case that RU is always in middle, Figure 9 is a comparison in case that RU is always in low, and Figure 10 is a comparison in case that RU is always in random numbers.

From analysis of Figure 7 in case that RU is always in high, our approach could achieve the user requirement with higher frequency than User-request approaches. In case that user requirement is higher than the service environment, it can be understood that the requirement cannot be fulfilled even if only user context is considered. Moreover, in case that agent considers only the user request and the maximum value that can be selected, URA generally is lower than our approach, when user requirement is not fulfilled. It is understood that some conflict on resource context is occurred. Our approach also decreases bad service construction by considering IAR. This is because that IAR decreases the conflict of resource context. From these results our approach is rather effective in this kind of case.

From analysis of Figure 8 in case that RU is always in middle, URA of User-request approach often closes to zero much more times than other approaches. On the other hand, compare to the User-request approach, IAR-based approach could fulfill the user requirement frequently. Moreover the case that URA of our approach is higher than that of Maximum approach is a little. From this result, if resources are available, our approach can provide a slightly better service than the original user requirement. We can find our approach considerably reduce bad service construction than other approaches such as in Figure 7.

By analyzing Figure 9 where RU is always in low, URA of User-request approach often extremely close to zero more

Figure 7: Results of comparison in case that RU is always in high
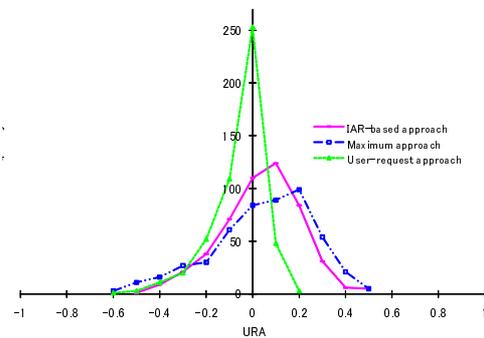
Figure 8: Results of comparison in case that RU is always in middle

frequently than other approaches. In our approach and Maximum approach, the case that URA closes to zero is not frequent. But, our approach closes to zero much more times than Maximum approach. Moreover User-request approach and our approach reach much closer to zero than Maximum approach. We can find that the agents construct organization considering user requirement and IAR effectively. However, our approach is thought to be meddlesome service for user who does not want excessive quality.

By analyzing Figure 10 where RU is always in random numbers, URA of User-request approach extremely close to zero more frequently than other approaches. In our approach, URA closes to zero and 0.2. Our approach also closes to zero much more times than Maximum approach. In negative side of URA, User-request approach generally is lower than our approach. Moreover, in Maximum approach, the case that URA closes to zero is not frequent. We can find that our approach that the agents try to construct organization to fulfill user requirement while avoiding low URA possible.

From these simulation results, it is understood that our approach is most effective under unstable environment with high-level user requirement. Additionally, it should be considered whether user requirement or relationship between agents which has the higher priority, when agents construct service. User requirement can be fulfilled by mainly considering IAR rather than user requirement in the environment with high

Figure 6: Behavioral situation representation of the system
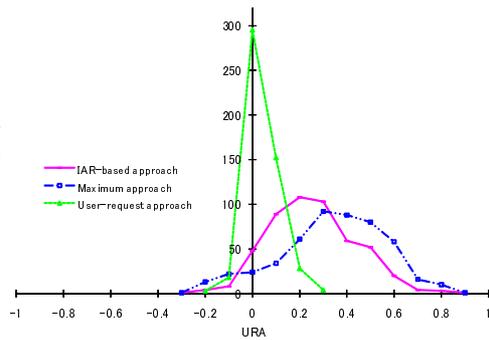
Figure 9: Results of comparison in case that RU is always in low
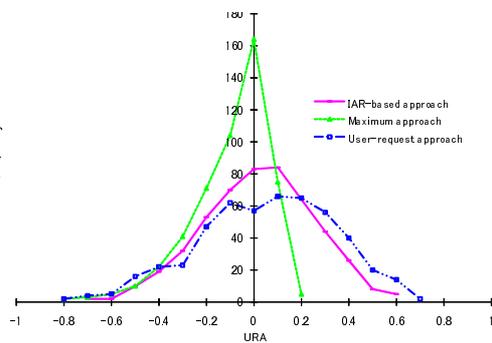
Figure 10: Results of comparison in case that RU is always in random numbers

user requirement and unstable resources. However, in the environment with suitable user requirement and stable resource, user requirement should be mainly considered rather than IAR. Therefore, we suppose that it is necessary to consider top priority between user requirement and IAR so that it matches to the situation when service is constructed in smart home environment.

6 APPLICATION

We developed two applications based on our middleware to confirm that the middleware can apply to a broad range of smart home services. One is a music distribution service and the other is a jam session service. These applications are in the entertainment application domain.

The first example is a music distribution service as shown in Figure 11. This service plays music from different speakers, by following the user's movement. As for hardware configurations for sensing the location information, we use an active-type RFID system. We put RFID receivers behind each speaker.

In Figure 12, a user is carrying a user terminal that is playing music with a tag of RFID. When the user approaches to the speaker with the user terminal, the music migrates from the user terminal to the nearest speaker. Speaker agents provided the service considering CR of IAR and user location.
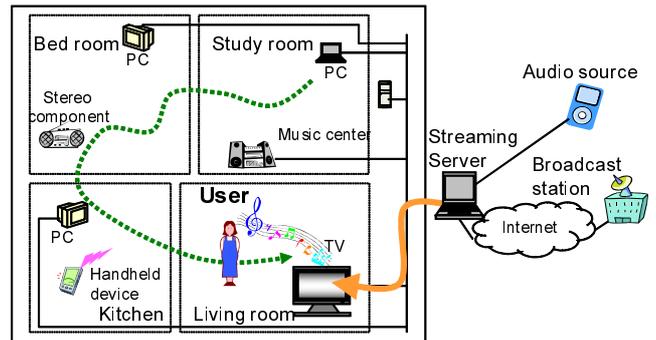
Figure 11: Overview of music distribution service
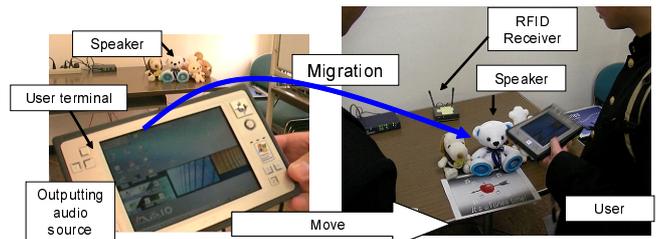
Figure 12: Snapshot of music distribution service

The second example is a jam session service as shown in Figure 13. It is an entertaining and tangible application that can be enjoyable for the people in the range from the children to the elderly. A user selects some paper cups on which RFID tag attaches. On the cup, a picture that represents an instrument is shown. When the user closes the cups to the speakers, the speaker plays instrumental audio sources corresponding to the paper cup's instrument picture, such as dram, piano, guitar, harmonica and synthesizer, via the Internet as shown Figure 14. If the user moves the paper cups away from the speaker, the speaker stops playing instrumental audio sources corresponding to the paper cup's instrumental picture. Therefore, the user can play instruments by combining some paper cups. This system's agent organization was constructed by RFID agents, Comp agents, AudioPlayer agents, Speaker agents, etc. In fact, we used two speakers, two RFID readers, and six RFID tags; this service was tried by more than 90 visitors at our laboratory's open house; the system continued to provide service without experiencing any problems.

We confirmed feasibility and effectiveness of our middle-
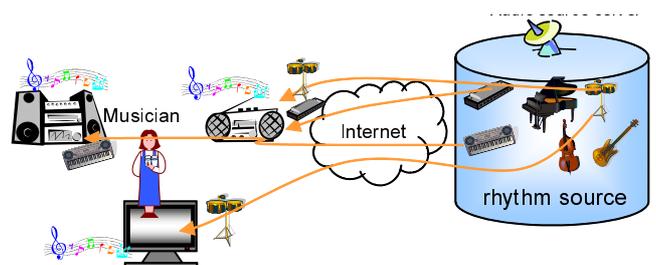
Figure 13: Overview of jam session service
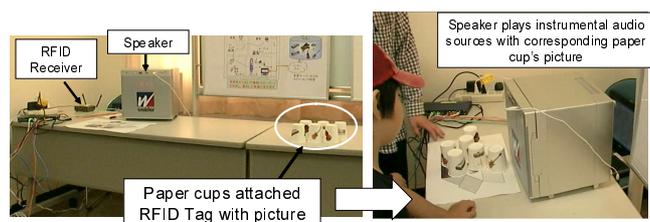
Figure 14: Snapshot of Jam session service

ware based on IAR through two applications.

7 CONCLUSION

In this paper, we described the concept of an agent-based middleware for smart home environment. We designed our middleware focusing on the service construction scheme for QoE-aware and energy-aware service provisioning considering the multiple contexts. We also evaluated our scheme with some simulation experiments and confirmed its usefulness in smart home environment, particularly for multimedia services. Moreover we implemented an application as a first step towards practical use of our middleware and performed some empirical studies with prototype system concentrating on evaluation of the effectiveness.

In future, we would like to design detail method and algorithm of policy-based energy management according to various situation. We are also planning to consider data fusion mechanism and evaluate prototype system using smart meter.

ACKNOWLEDGEMENT

REFERENCES

[1] M. Weiser, The Computer for the Twenty-first Century, Scientific American, Vol. 265, No. 3, pp. 94–104 (1991).

[2] H.-M. Kim, and T. Kinoshita, A Multiagent System for Microgrid Operation in the Grid-interconnected Mode, Journal of Electrical Engineering & Technologys, Vol. 5, No. 2, pp. 246–254 (2010).

[3] R. Huan, M. Itou, T. Tamura, and J. Ma, Agents based Approach for Smart Eco-home Environments, Proc. of 2010 IEEE World Congress on Computational Intelligence (WCCI 2010), pp. 1–8 (2010).

[4] N. Shiratori, K. Hashimoto, D. Chakraborty, H. Takahashi, T. Suganuma, N. Nakamura, and A. Takeda, Kurihara Green ICT Project – Towards Symbiosis between Human's Life and Nature, Journal of Internet Technology(JIT), Vol. 12, No. 1, pp. 1–11 (2011).

[5] T. Itao, T. Nakamura, M. Matsuo, T. Suda, and T. Aoyama, Adaptive Creation of Network Applications in the Jack-in-the-Net Architecture, Proc. of IFIP Networking, pp. 129–140 (2002).

[6] M. Minoh and T. Kamae, Networked Appliance and Their Peer-to-Peer Architecture AMIDEN, IEEE Comm. Magazine, vol. 39, No. 10, pp. 80–84, (2001).

[7] T. Iwamoto and H. Tokuda, PMAA:Media Access Architecture for Ubiquitous Computing, Journal of IPSJ, Vol. 44, No. 3, pp. 848–856 (2003).

[8] M. Minami, H. Morikawa, and T. Aoyama, The Design and Evaluation of an Interface-based Naming System for Supporting Service Synthesis in Ubiquitous Computing Environment, Joutnal of IEICEJ, Vol. J86-B, No. 5, pp. 777–789 (2003).

[9] J. Nakazawa, Y. Tobe, and H. Tokuda, On Dynamic Service Integration in VNA Architecture, IEICE Trans., Vol. E84-A, No. 7, pp. 1624–1635, (2001).

[10] S. Gribble, M. Welsh, R. Behren, E. Brewer, D. Culler, N. Borisov, S. Czerwinski, R. Gummadi, J. Hill, A. Joseph, R. Katz, Z. Mao, S. Ross, and B. Zhao, The Ninja architecture for robust Internet-scale systems and services, Special Issue of Computer Networks on Pervasive Computing, Vol. 35, No. 4 (2001).

[11] H. Takahashi, T. Suganuma, and N. Shiratori, AMUSE: An Agent-based Middleware for Context-aware Ubiquitous Services, Proc. of the International Conference on Parallel and Distributed Systems (ICPADS2005), pp. 743–749 (2005).

[12] Y. Sato, H. Takahashi, T. Suganuma, and N. Shiratori, Design and Implementation of an Adaptive Control Scheme for Context Information Delivery, Proc. of the 3rd International Conference on Complex, Intelligent and Software Intensive Systems (CISIS-2009), pp. 351–358 (2009).

[13] P. Bellavista, A. Corradi, R. Montanari, and C. Stefanelli, Context-Aware Middleware for Resource Management in the Wireless Internet, IEEE Trans. software Eng., Vol. 29, No. 12, pp. 1086-1099 (2003).

[14] W. Y. Lum and F. C. M. Lau, User-Centric Content Negotiation for Effective Adaptation Service in Mobile Computing, IEEE Trans. software Eng., Vol. 29, No. 12, pp. 1100–1111 (2003).

[15] J. P. Sousa, V. Poladian, D. Garlan, B. Schmerl, and Mary Shaw, Task-based Adaptation for Ubiquitous Computing, IEEE Trans. Systems, Man, and Cybernetics, Part C: Applications and Reviews, Vol. 36, No. 3, pp. 328–340 (2006).

[16] R. G. Smith, The Contract net protocol: High-level communication and control in a distributed Problem solver, IEEE Trans. Comput., Vol. 29, No. 12, pp. 1104–1113 (1980).

[17] K. Sugawara, H. Hara, T. Kinoshita, and T. Uchiya, Flexible Distributed Agent System programmed by a Rule-based Language, Proc. of the Sixth IASTED International Conference of Artificial Intelligence and Soft Computing, pp. 7–12, (2002).

[18] T. Uchiya, T. Maemura, H. Hara, K. Sugawara, and T. Kinoshita, Interactive Design Method of Agent System for Symbiotic Computing, International Journal of Cognitive Informatics and Natural Intelligence, Vol. 3, No. 1, pp. 57–74 (2008).

A Study on Office Environmental Control System Using Wireless Sensor Networks

Xingang HE†, Tadanori MIZUNO*, and Hiroshi MINENO††

†Graduate School of Informatics, Shizuoka University, Japan
* Faculty of Information Science, Aichi Institute of Technology, Japan
††Faculty of Informatics, Shizuoka University, Japan
{xin, mineno}@minelab.jp, mizuno@mizulab.net

Abstract - We propose an office environmental control system using wireless sensor networks (WSNs) to prevent power wastage. This system collects sensing data such as temperature, motion, humidity, luminance, and electric power from offices or families using environmental sensor nodes and smart sockets. By analyzing these sensing data, the system is able to control different kinds of electrical appliances automatically. We tried to use our system to control monitors and fluorescent lights in our laboratory during two experiments. We compared the power consumption with and without the system control. The experimental results confirmed that the system was able to reduce power consumption by about 20%.

Keywords: environmental control, wireless sensor network, energy conservation

1   INTRODUCTION

Energy consumption has recently been increasing being influenced by changes in the national lifestyle by people pursuing and comfort. According to *Energy in Japan 2010* [1] announced by the Agency for Natural Resources and Energy in Japan, energy consumption is growing rapidly in commercial/residential sectors, in which we directly use energy at home or in the office. Since energy has been overused and can contribute to global warming, energy conservation has become a worldwide problem that everyone has to be conscious of. After the unprecedented earthquake and the nuclear crisis happened in March, 2011 in Japan, reducing the amount of electrical energy consumed has become more important than ever before. Home energy management systems (HEMS) and building energy management systems (BEMS) have recently been proposed to conserve energy while retaining the same convenience and comfort afforded by appliances [2]. The main idea behind these systems is to automatically or indirectly control appliances at homes/offices while providing more convenient lifestyles and reducing the energy consumption at the same time.

Wireless sensor networks (WSNs) are widely used for many different purposes like farm management [3], positioning [3], [4], security management [5], [6], industry, and medical science because of their advantageous characteristics such as low energy consumption, ease of use, and dynamic network topology.

This paper describes an office environmental control system using WSNs. It consists of several environment sensor nodes, smart sockets, infrared remote control nodes, a WebIO [7], a sensor gateway, and a database. First, we collect sensing data such as those on temperature, motion, humidity, and luminance using environment sensor nodes that are installed in all seats or hung from ceilings to understand information from offices. We simultaneously use smart sockets to measure the electric power of all appliances to monitor their energy consumption in rooms. We then use a sensor gateway to analyze received sensing data, using them to control the existing appliances through infrared remote control nodes or WebIO. In the following text, section 2 explains related work and section 3 presents details on the structure and components of our office environmental control system. Section 4 describes the experiments and discusses our evaluation of the effect of power saving achieved by using our system. Section 5 is the conclusion and discusses future work.

2   RELATED WORK

Much work has been done on energy conservation systems by using WSNs because they can be used to measure different situations in rooms such as temperature and motion. *An Energy Conservation System for Intelligent Buildings by Wireless Sensor Networks* [8] has been proposed. The system detects whether electricity is possibly being wasted by WSNs and then turns these unnecessary electric appliances off via power-line control devices. However, since power-line control devices can only adjust the electrical currents of appliances, the system can only have limited control such as turning off appliances. Users utilizing the system also do not know how much electric power the appliances are using. Therefore, it is difficult to evaluate the power consumed in rooms.

Several major companies in the electrical appliance industry have recently been trying to develop HEMS systems to solve these problems by implementing the concept of smart houses, which cooperate with all appliances at home [9], [10]. The concept of smart houses is supposed to satisfy several requirements, such as managing power consumption in houses, automatically controlling appliances like lighting fixtures using kinds of sensors, and using clean energy like that from solar energy. However, even though smart houses have many advantages, a huge barrier is that users need to replace almost all their existing appliances with new ones to support the whole system, which involves enormous cost.

Compared with related work, we can use an infrared remote control node in our system to control existing appliances that can be operated by infrared remote control. The system can also monitor the electric power consumed by appliances in real time because the smart sockets measure the electric power of appliances. This means users can be aware of how

much electric power is used by existing appliances in rooms. Therefore, it is also much easier to evaluate whether the system is working well or not.

3 OFFICE ENVIRONMENTAL CONTROL SYSTEM

3.1 Overview

We developed a prototype of the office environmental control system and **Fig. 1** shows an overview of it. First, the system uses a parent node to collect different kinds of sensing data such as motion, illuminance, temperature, and power consumed by appliances from environment sensor nodes (ESNs), and smart sockets that are installed at every seat. The sensor gateway stores the sensing data into a database and analyzes them. After it has analyzed the sensing data, the sensor gateway determines what control action should be taken and generates appropriate control commands. The sensor gateway then sends the control commands to the controlling devices. We had two ways of controlling devices in our system. The first was by using infrared remote control nodes (IRCNs) to control the appliances that could be controlled by infrared remote control; the second way was by using WebIO to control other appliances like fluorescent lights that could not be controlled by infrared remote control. We also developed an application that runs on Android smart phones. Users can check the status of appliances, manage the sensor gateway system, or control appliances directly by using smart phones installed with applications.

3.2 Components

Our system uses seven kinds of components. These components include environment sensor nodes (parent and children), smart sockets, a sensor gateway, a database, infrared remote control nodes, WebIO, and smart phones. These components are introduced as follows.

3.2.1 Environment sensor nodes (parent and children)

The environment sensor nodes (ESNs) were designed by Shizuoka University [11]. They involve four kinds of sensors: infrared motion, illuminance, temperature and humidity. Each ESN also contains a wireless communication module compliant with the IEEE 802.15.4 standard.

The ESNs play two roles of a parent or children. Users can change their roles easily with a small switch. A parent node is like a sink node in most WSNs. It collects all the sensing data from all child nodes and smart sockets that are sent to the sensor gateway by serial communication. The sensor gateway uses the parent node to send control commands to IRCNs. Child nodes send all kinds of sensing data to the parent node immediately if any motion is detected. However, if no motion is detected, child sensors send sensing data to the parent node every 10 minute. We assumed there was only one parent ESN and several child ESNs in our system.
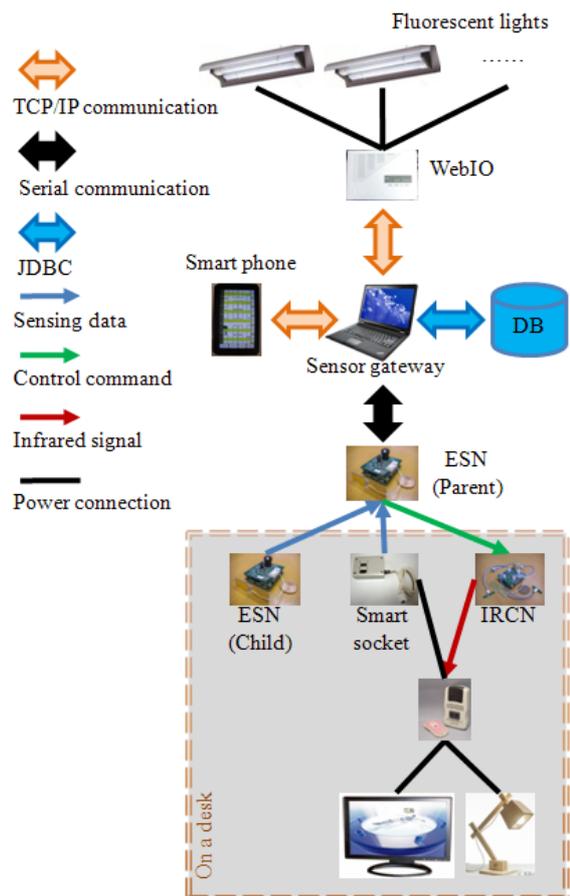
Figure 1: Overview of office environmental control system

3.2.2 Smart sockets

The smart sockets were designed by Shizuoka University. They included a sensor that could measure the electric power of appliances connected to it. Like ESNs, all smart sockets also contain a wireless communication module compliant with the IEEE 802.15.4 standard. Smart sockets are supplied with electricity by being connected to electric outlets. Thus, if any appliances are connected to smart sockets, the sockets will supply them with electricity. The electric power of appliances in real time is simultaneously measured and sent to the parent node every second.

3.2.3 Sensor gateway

The sensor gateway is a major part of the system and we used a computer for this. It plays 4 roles: First, it obtains all the sensing data from the parent node by serial communication. Second, it analyzes the sensing data, and then generates related control commands. Third, it sends the control commands to IRCNs using the parent node or to WebIO by TCP/IP communication. Fourth, it sends all sensing and log data to the database. Moreover, the sensor gateway can be connected to a smart phone. The sensor gateway can send the statuses of appliances to smart phones and turn automatic control on or off on some appliances according to the smart phone settings.

3.2.4 Database

The database stores all the sensing data and controls the log data from the sensor gateway. The sensor gateway will sometimes use some of these data to generate control commands. These data can also be accessed by other systems such as sensing data visualization systems, which display power consumption.

3.2.5 Infrared remote control nodes

The infrared remote control nodes (IRCNs) are devices that remotely control appliances with infrared signals. These were also designed by Shizuoka University. The IRCNs were implemented by using an infrared signal learning device and an infrared sending device. An IRCN can learn 20 kinds of infrared signals and save them to an electronically erasable programmable read-only memory (EEPROM) by using the infrared signal learning device. With infrared sending device, on the other hand, an IRCN can only send infrared signals that it has learned previously. Each IRCN also contains a wireless communication module compliant with the IEEE 802.15.4 standard like ESN and smart sockets. The sensor gateway can use the parent node to control IRCNs to learn any infrared signals or send infrared signals that have been learned to appliances.

3.2.6 WebIO

WebIO is a device that controls appliances such as fluorescent lights without infrared remote control. It is connected to sensor networks by LAN. The sensor gateway can control one or more appliances at a time by only sending one control command to WebIO at one time using TCP/IP communication. Since there are 16 output ends on a WebIO, the sensor gateway can use one WebIO to control 16 different appliances at the same time. Because WebIO stores the status of each output end, the sensor network can send a required command to WebIO to obtain the status information of each appliance being connected.

3.2.7 Smart phones

We developed an application that runs on Android smart phones. After our application is installed, smart phones can connect to the sensor gateway by TCP/IP communication. Users can check the states of appliances remotely with smart phones. They can also use smart phones to turn the automatic control on or off, or control the application directly by using smart phones.

4 EXPERIMENTS AND EVALUATION

This section describes our experiments using the office environmental control system, which we evaluated by using the experiment results.

We carried out our experiments in a room at our laboratory that was about 80 square meters in area. There were 14 students in our room and they all had their own seat. **Table 1** lists the devices we used in the experiment. We used one

Table 1: Devices in experiments

Device	Number
Sensor gateway	1
Database	1
ESN (parent)	1
ESNs (children)	14 (seat) + 18 (ceiling) = 32
IRCNs	14
Smart sockets	14
WebIO	1

sensor gateway that was connected to one ESN (parent) to collect sensing data from all the ESNs (children) and smart sockets. There were 32 ESNs (children) used to collect environmental sensing data. Fourteen of them were installed, i.e, one for each seat. The other 18 were hung from the ceiling to collect motion sensing data from the aisles. **Fig. 5** outlines the platform for room in the experiments and the position of the ESNs. We installed a smart socket in every seat that connected all the appliances to the same seat to collect data on the electric power consumed by each student. We could therefore monitor the electric power consumed by every seat (student). We also installed an IRCN in every seat to turn the monitors on and off. There were three devices in every seat: an ESN (child), a smart socket, and an IRCN. Moreover, to control all of the fluorescent lights in experiment room, we installed a WebIO in the room for the experiment and it was connected to the sensor gateway using a LAN.

We carried out two experiments to evaluate the effect of conserving power consumption. The first was using the system to control the appliances in each student's seat. The second was using the system to control fluorescent lights. We will describe details on both experiments in the following text.

4.1 Experiment to control monitors

We use the system to conduct an experiment to control monitors. Since we had installed an ESN (child) and a smart socket in every seat, we could monitor the results for motion and electric power over time for every seat. **Fig. 2** plots the results for electric power and detected motion per day for two seats (A and B) without the system control. The red lines plot the electric power for a seat at that moment, while the blue marks indicate the periods motion was detected. Fig. 2(a) plots the electric power for seat A that was maintained at about 180 W all through the day. However, from the result for detected motion, we surmised that the student on that seat did not sit on that seat all through the day, because motion was only intermittently detected from 10:00 to 21:00. Compared with seat A, Fig. 2(b) shows the electric power of seat B changing over time. Each time motion was not detected for about 30 minutes, the electric power reduced from 180 to 110 W.

We asked the students at these two seats why the differences had occurred. The main reason was that they applied different settings to power management on the OS. The student in seat A only used the screen saver when there was no response from the mouse or keyboard after a few minutes.
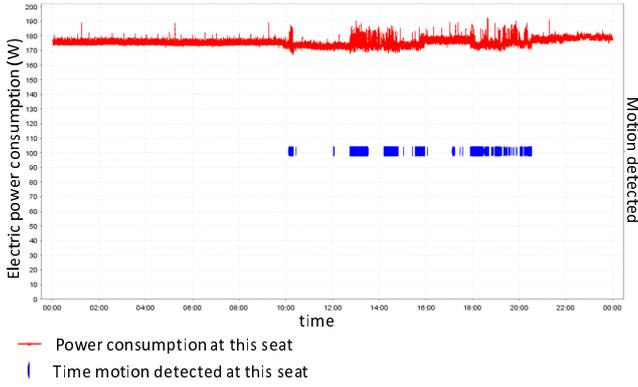
(a) Seat A
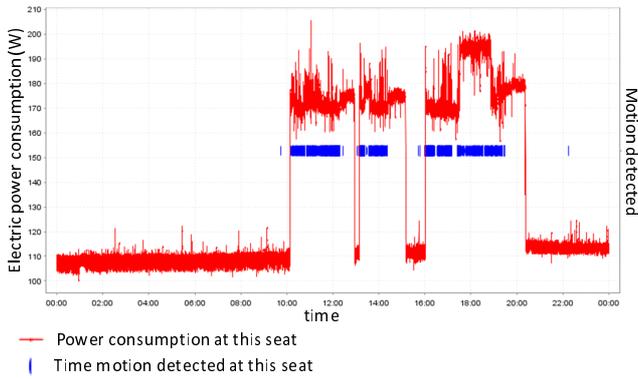
(b) Seat B

Figure 2: Results for electric power and detected motion without system control per day

Figure 3: Devices used for controlling monitors
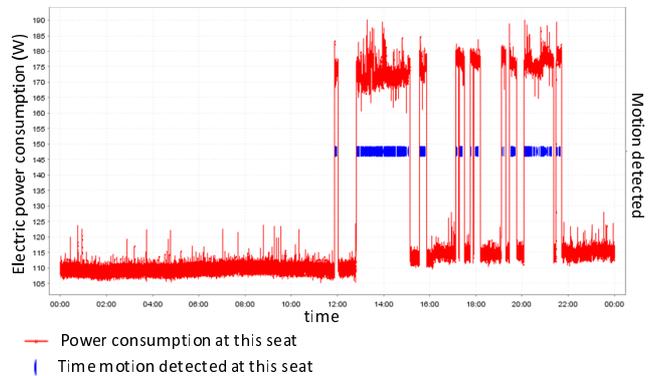
Figure 4: Results for electric power and detected motion with system control per day

However, the student in seat B applied power management on the OS to turn monitors to standby status if there was no response from the mouse or keyboard after 30 minutes.

Most of our students needed to keep their desktops on for a long time because of research requirements. However, appliances like monitors do not need to be kept on when students are absent. Moreover, all the students in our laboratory used more than two monitors. That meant that electric power consumed by monitors was more than 70 W per seat, which was not a small amount of wattage. Therefore, we tried to use our system to control monitors automatically to reduce unnecessary power consumption. **Fig. 3** has a photograph of the experiment devices we installed on a seat. Two monitors on the desk were connected to an infrared remote control socket. Therefore, our system could use an IRCN to control monitors by controlling the infrared remote control socket.

We adjusted the system as follows: if no motion was detected by the ESN in this seat for more than 5 minutes and the monitors were in on status, the system turned them off by using IRCN. And if any motion was detected by the ESN in this seat and the monitors were in off status, the system turned them on by using IRCN. **Fig. 4** plots the results for a seat with the system control. We can see that the electric power reached a high level only when motion was detected, and after no motion was detected for 5 minutes, the electric power of the seat

returned to a low level.

We tried to estimate the energy saved, the rate at which the monitors' power consumption was reduced by using our system with the following equation:

$$(((W \times (x-p) \times n/60)) / (W \times t)) * 100\% ,$$

where "W" is the energy used by monitors, "x" is the time to turn monitors to standby mode, which is set in power management on the OS, and "p" is the time to turn off the monitors, which is set in the sensor gateway. Here, "n" is the number of times that a student in the seat left his seat for more than "x" minutes, and "t" is the time that a student at the laboratory per day.

We asked all the students at our laboratory about their habits and calculated the power consumption that was reduced by using our system. Most students set the "x" to 30 minutes. We set "p" at the gateway to 5 minutes. Because most students left their seats about five times a day for lunch, classes or meetings, we set the "n" to five times. We also set "t" to 10 hours according to the results from the time cards. We calculated that in this case we could use our system to reduce the power consumed by the monitors by about 20%.
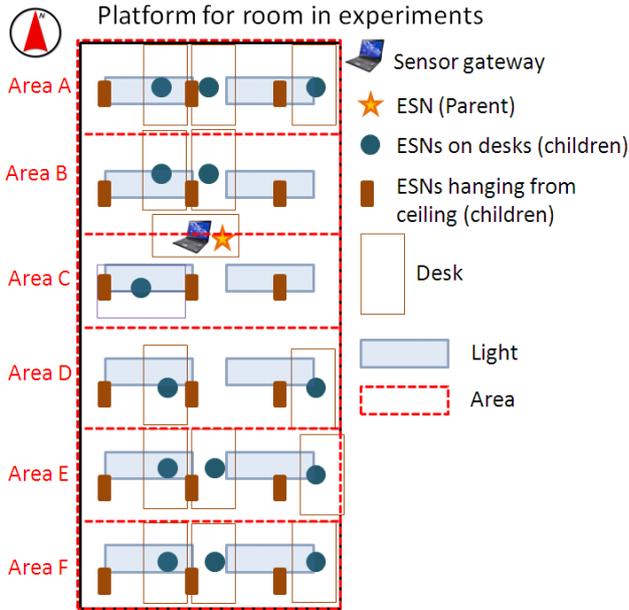
Figure 5: Platform for room in experiments
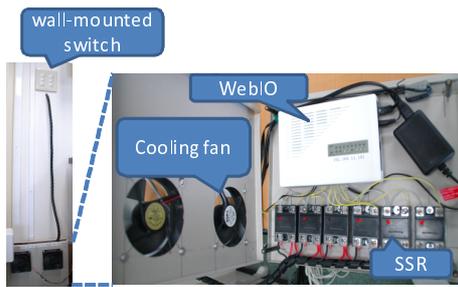
Figure 6: Devices used to control fluorescent lights

4.2 Experiment to control fluorescent lights

We used our system to conduct another experiment that controlled the fluorescent lights in our laboratory. **Fig. 5** outlines the platform for the room at our laboratory. The positions of the ESNs and lights are in the figure. There are 12 fluorescent lights at our laboratory. The room is divided into six areas by the six red dashed frames. The lights in the same area are controlled by the same wall-mounted switch. In this experiment, we assumed the sensing data of the ESNs were in the same area to control the lights that were in the same area. However, because fluorescent lights can not be controlled by infrared remote control, we have to do further work on controlling lights with our system. **Fig. 6** is a photograph of the devices we used to control the lights. We ran six cables out from the wall-mounted switch, and connected them to six Solid-state relays (SSRs). We placed the WebIO and six SSRs in a box, with two cooling fans to help release heat. **Fig. 7** has circuit diagrams for one light. We can see that light could only be controlled by the switch before. However, with the SSR and WebIO, the light can also be controlled by the WebIO. For example, if the WebIO receives a command from the sensor gateway to turn on one light, the WebIO will turn on
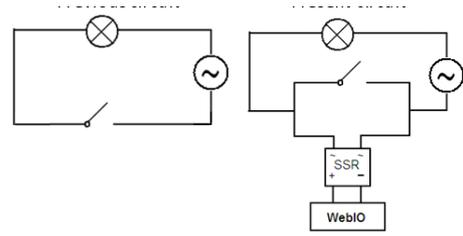
Figure 7: Circuits for one light
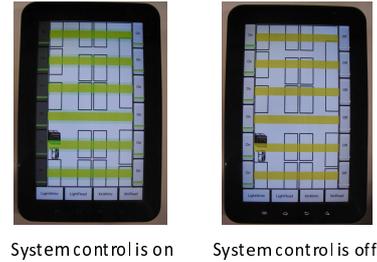
Figure 8: Interface for application running on Android smart phone connected to control system

the SSR, so that the light will be turned on at the same time.

We set the system similarly to that to control the monitors: if no motion was detected by any ESN in the same area for more than 5 minutes and the lights in this area were in on status, system turned off the lights in the area using WebIO. If any motion by any ESN in the same area was detected and lights in this area were in off status, the system turned the lights on in the area using WebIO.
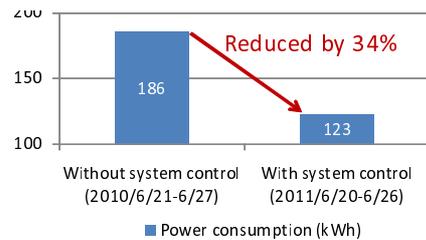
(a) Compared with same period last year
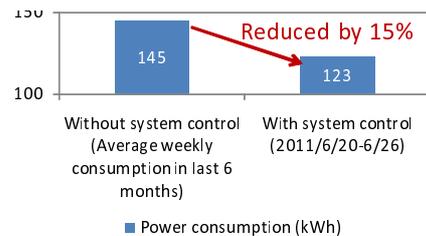
(b) Compared with average weekly consumption in last 6 months

Figure 9: Comparison of power consumption

Moreover, an Android smart phone can be used to check the status of lights and cancel or restart the automatic system control on selected lights. **Fig. 8** has photographs of interface for an Android phone.

We used our system to control lights automatically for one week (2011/6/20 - 2011/6/26). Then, we compared the power consumed by the lights to that for the same period last year (2010/6/21 - 2010/6/27). (Fig. 9(a)) Compared to the same period last year, using our system to control lights reduced by about 34% power consumption. Even when compared to the average weekly power consumption for the last six months (removing months in which the most and least power was consumed), our system reduced power consumption by about 15%. (Fig. 9(b))

5 CONCLUSION

We proposed an office environmental control system using wireless sensor networks (WSNs) to prevent power wastage. This system collected sensing data from all the areas of an office or family room by using environment sensor nodes and smart sockets. The system was able to control electrical appliances automatically by analyzing the sensing data. We used our control system in the experiments to manage appliances such as monitors and fluorescent lights. The results from the experiments confirmed that power consumption could be reduced efficiently by using our system. However, there is one limitation in that if people remain motionless for longer than the setting time of the system, it will turn appliances off. In the future, we intend to aim at preventing the erroneous control caused by current limitations and increasing the comfort level, using different controlling intervals of time depending on different users' habits. Moreover, we intend to use other kinds of sensing data to control other appliances such as air conditioning.

REFERENCES

[1] Agency for Natural Resources and Energy, Japan. "the Energy in Japan 2010", 2010.

[2] H. Mineno, Y. Kato, K. Obata, H. Kuriyama, K. Abe, N. Ishikawa, T. Mizuno "Adaptive Home/Building Energy Management System Using Heterogeneous Sensor/Actuator Networks", in CCNC'10, Jan. 9-12 2010.

[3] Di Wu; Renfa Li; Ruifang Li; Hualin Liang; Juan Luo Localization in UWB Sensor Networks for Agriculture Environment, in *Wireless Communications, Networking and Mobile Computing, 2008. WiCOM '08. 4th International Conference*, Oct. 2008.

[4] Sung-Hwa Hong; Byoung-Kug Kim; Doo-Seop Eom Localization algorithm in wireless sensor networks with network mobility, in *Consumer Electronics, 2009 IEEE International Conference*, Nov. 2009.

[5] Ying-Wen Bai, Li-Sih Shen and Zong-Han Li Design and Implementation of an Embedded Home Surveillance System by Use of Multiple Ultrasonic Sensors, in *Consumer Electronics (ICCE), 2011 IEEE International Conference*, Jan. 2011.

[6] Zualkernan, I.A.; Al-Ali, A.R.; Jabbar, M.A.; Zabalawi, I.; Wasfy, A. InfoPods: Zigbee-based remote information monitoring devices for smart-homes, in *Consumer Electronics, IEEE Transactions*, Aug. 2009.

[7] http://www.lifetron.jp/products.htm

[8] Yu-Chee Tseng, You-Chiun Wang, and Lun-Wu Yeh iPower: an energy conservation system for intelligent buildings by wireless sensor networks, in *International Journal of Sensor Networks*, page(1 - 10) Volume 5, 2009.

[9] http://www.toshiba.co.jp/env/jp/
energy/index_j.htm

[10] http://www.sharp.co.jp/corporate/rd/
36/pdf/101_04.pdf

[11] Hiroshi Mineno, Keiichi Abe and Tadanori Mizuno An Adaptive Energy Management System Using Heterogeneous Sensor/Actuator Networks, in *InTech, August 2011*, August 2011.

A Method for Modeling of Pedestrian Flow in the Space with Obstacles using Laser Range Scanners

Yoshitaka Nakamura[†], Yusuke Wada[‡], Teruo Higashino[‡], and Osamu Takahasi[†]

[†]Faculty of Systems Information Science, Future University Hakodate, Japan
[‡]Graduate School of Information Science and Technology, Osaka University, Japan
{y-nakamr, osamu}@fun.ac.jp, {w-yuuki, higashino}@ist.osaka-u.ac.jp

Abstract - The measurement of the pedestrians' movement has high utility, because many efficient services can be provided for pedestrians with the measurement data. There are some techniques to measure pedestrian traffic and especially the method to measure pedestrian traffic with laser range scanners attracts considerable attention recently. However, the lines of sights of the laser range scanners are obscured by obstacles. Therefore, it is difficult to completely capture all the pedestrians' movement in the target area. In this paper, we propose a method that estimates pedestrian traffic from sectional population density instead of individual pedestrians' data. Our method also generates pedestrians flow. We have evaluated the proposed method with ideal scenario data. The experimental results have shown that our method could reproduce pedestrian traffic with about 80% accuracy.

Keywords: Laser Range Scanner, Estimation of Trajectory, Pedestrian Flow

1 INTRODUCTION

With the recent development of the sensing technology, various phenomena can measured in the real world and the services using these measurement data become able to be provided. As such an example, Cyber Physical System attracts attention. The construction of the smart society by various services using various sensors becomes one of the new issues in the future information and communication technology. Among phenomena to be targeted for the measurement of the sensor, a tendency of the movement of the pedestrians (pedestrian flow) attracts attention recently. The information of pedestrian flow becomes important because this information in roads or the underground shopping center is available for the orientation of various services and the control of pedestrian flow.

For the mainstream method of the pedestrian flow measurement, there is the method that processes the images recorded by cameras, and tracks the pedestrian using the information. However, in the method using the camera, there are the privacy problem of target pedestrians and the problem of setting cost and the angle of view of cameras. In contrast, the measurement of the pedestrian flow with the laser range scanner attracts attention from a demand to measure the rough tendency of the pedestrian flow at low cost. The laser range scanner can scan a wide area fast. There is little fear to infringe the privacy information of pedestrian targeted for a measurement, because the measurement data which laser range scanner acquires are only positional information expressed as direction

and distance for the scanner. In addition, the calculation cost is relatively small because the size of scanner data is much smaller than images. However, the laser range scanner is easy to lose sight of measurement objects by occlusion, and it is difficult to completely measure the all behaviors of all pedestrians in the measurement area.

In this paper, we propose the method for modeling pedestrian flow by estimation of the routes that the pedestrians would have passed not from the individual positional data of each pedestrian but from the population density at each division of the measurement area. In our method, the measurement area is divided in the division of constant size at first. Next, in each division, the number of pedestrians is counted using laser range scanner, and the population density in the division is calculated. Finally the course where pedestrian passed is estimated from the population density and generate pedestrian flow model. In this way, our method can generate the pedestrian flow model from the data that the all behaviors of all pedestrians are not observed.

2 RELATED WORK

2.1 Methods of measuring pedestrian flow

There are various methods to track the pedestrians. For example, a tracking method using images recorded by video cameras is often used. There is a method to measure pedestrian flow by recognizing the head of the pedestrians and tracking automatically using a stereovision camera. However, the measurement using cameras has much quantity of data, and the calculation cost becomes high. And from the viewpoint of privacy, the setting of the camera may be restricted.

There is another method to let each pedestrian hold the RFID tag which sends unique ID. This method estimates the position of each pedestrian from the position of the ID receiver and ID sent from RFID tags, and tracks pedestrians. However, for the tracking using the RFID tag, it is necessary to build the environment where ID receivers are deployed enough and to let pedestrians hold the terminal with RFID tag.

There is a method to estimate overall pedestrian flows from the flow quantity of each gateway by counting the passage number of people. However, if the instrumentations are affected by the occlusions, there is a problem that the numbers of the pedestrians who are not measured definitely increase as quantity of traffic increases. In addition, because only the number of passing pedestrians is measured, some behaviors such as stopping in the measurement area cannot be detected.

Some method using laser range scanners are proposed. One

of them measures the ankle of the pedestrian using scanners, detects walking rhythm of the bipedalism from the movement pattern of the ankle and tracks the pedestrian. However, in actual environment, this method may be influenced by the existence of the bag with the caster. Therefore it is necessary to devise measurement procedure.

2.2 Method of generating pedestrian flow model

The generation of pedestrian flow model is mainly used in the simulation of mobile wireless networks. Simple models such as Random Way Point Model[1] are often used for pseudo node mobility models. In late years some researches[2]–[5] to propose the mobility equal to subspecies of Random Way Point Model are performed. Many mobility models to compose realistic mobility by the measurement data and the geographical information have been proposed. Ref.[6] propose the model to reproduce the interception of the radio wave by the buildings(obstacles), and the mobility model that nodes avoid those buildings. The method of Ref.[7] divides the simulation domain into some zones depending on the characteristic such as a residential area or a business district, and estimates the change of the density of every node classification by zonal unit using existing traffic planning method. In Ref.[8], Weighted Way Point Model is proposed. This method defines the domains with many people such as a cafe or the university, and models the movement of nodes between domains using Markov model by giving distribution of the sojourn time in each domain and the transition probability between domains. Time Slot Urban Pedestrian Flows model[10] estimates realistic movement routes on given road structure and derives the traffic quantity of each route to satisfy given density.

3 PEDESTRIAN FLOW MEASUREMENT USING LASER RANGE SCANNERS

3.1 Characteristic of the laser range scanner

A laser range scanner is the sensor which can measure the distance from a sensor to an object using the spread time of laser beam. And this sensor has characteristic to scan wide area fast. But, there is the fault to lose sight of the pedestrian targeted for the measurement when the pedestrian is hidden by the shade of obstacles such as pillars or different pedestrians(the occlusion problem). The laser range scanners which we used for pedestrian flow measurement is UTM-30LX[12] made by HOKUYO AUTOMATIC CO., LTD. Table1 shows the specifications of this sensor. As for this sensor, the tracking of a pedestrian targeted for the measurement is possible[17].

3.2 Precedent experiment

A precedent measurement experiment was conducted in underground shopping center "Whity Umeda" of the neighborhood of subway Umeda Station of Osaka. We measured the pedestrians by synchronizing four UTM-30LXs. Fig.1 shows the measurement area and the installing position of the laser range scanners. The installed laser range scanners are

Table 1: UTM-30LX

Item	Spec
Detection Range	$0.1 - 30m, 270°$
Angular Resolution	$0.25°$
Scan Time	$25ms/scan$
Accuracy	$0.1 - 10m : \pm30mm$
	$10 - 30m : \pm50mm$
Size	$W60 \times D60 \times H87mm$
Weight	$370g$

expressed as the orange column of Fig.1, the movable range of pedestrian is expressed with beige. The data measured by
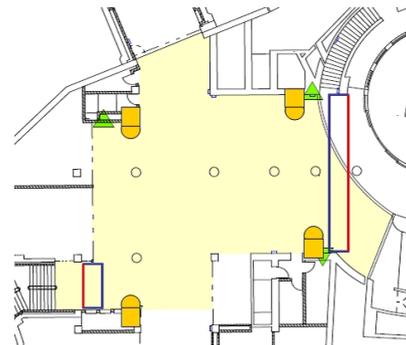
Figure 1: Installation Position of Laser Range Scanners

UTM-30LX are position coordinate data of the pedestrian in the measurement area every one second, and the unit of coordinate supports mm of the real world. In addition, specific ID is assigned to the pedestrian who appeared newly in the measurement area. While the tracking with the sensor succeeds, the identification of the same person is possible by this ID because the ID is unchangeable.

The data measured by UTM-30LX are following three.

- Measurement time: t

- Position coordinates of the pedestrian: (x, y)

- ID of each pedestrian assigned to by the tracking of the laser range scanners: i

After analyzing actual measurement data, the tracking in UTM-30LX is proved that it could continue only for a short time by the occlusion problem. The life time of the ID from 30% to 40% was approximately one second. It is caused by the following that a success period of the tracking shortens.

- A pedestrian is hidden behind obstacles such as pillars

- A pedestrian is hidden behind other pedestrians

- There are some pedestrian staying by the wall where sensors are installed and they narrow the measurement range of the sensor.

There are some methods to solve these problems. For example, they are methods to increase the number of sensors and method to use sensors which are not affected by obstacles. However, these methods have high cost.

4 PROPOSED METHOD

4.1 Overview

Because of the characteristic of laser range scanner, it is difficult to completely measure the behavior of all pedestrians when laser range scanners are used for measuring pedestrian flow. On the other hand, changes of pedestrian flow as the whole is regarded more important than each personal behavior in the scene where the pedestrian flow is used including trajectory analysis of customer in commercial facilities and pedestrian flow analysis for refuge instructions. In this paper, we pay attention to the change of the population density in the partial domain in the measurement area. Generally, the population density on the route that many people pass becomes higher, and population density changes along the direction of the pedestrian flow. Using this characteristic, our method extracts the characteristic of the pedestrian flow in the measurement area.

In our method, at first the measurement area is divided in the plural square domains of the equal size (unit cells), and the population density of each cell is calculated by the positional data of the pedestrian. From a tendency of the population density of each cell, our method estimates the cell which the pedestrians passed and estimates the routes of the pedestrians. The direction ratio of pedestrian via the estimated route is decided based on the directional information of pedestrian provided by the tracking with the laser range scanners. Our method expresses the number of pedestrians via certain route and the movement direction of the route in the form of the flow. This flow is the proposed pedestrian flow model.

4.2 Generation method of pedestrian flow model based on the population density

In this method, we use the measurement data and the tracking data of the laser range scanners for the measurement of the pedestrian. The tracking of the walker with the range scanner succeeds only for several seconds, 30% - 40% of IDs are measured only for about one second. Some pedestrians are expected to continue moving without being measured by sensors for a long period. Therefore to estimate the movement of the pedestrian we use the population density of the cells for certain constant period. The following assumptions are put for the movement of the pedestrian.

- Assumption 1 Pedestrians moves to the adjacent cell of top and bottom right and left from the cell which oneself is now.

- Assumption 2 Pedestrians move from the entrance to the exit without making a detour under Assumption 1.

- Assumption 3 Pedestrians are measured once in each cell which he/she passes.

The pedestrian flow model is generated in the following procedures.

1. Division to the unit cells of the measurement area

2. Choice of the gateway cells

3. Supposition of the route candidate between the gateways

4. Calculation of the population density of each cell

5. Calculation of the number of sojourners

6. Estimation of the route where the pedestrian passed and its traffic

7. Determination of direction of the flow

The details of each processing are as follows.

1. Division to the unit cells The measurement area is divided into cells like Fig.2. Any size can set as the size of the cell. But it becomes hard to get the tendency of the movement of the pedestrian group as much as the size becomes large. And it costs for the calculation of the movement route of the pedestrian as much as the size becomes small. We assume the size of the cell 6.3m every direction in consideration of the speed of the pedestrian and the size of gateways.

Figure 2: Target Area Divided into Unit Cell

2. Choice of the gateway cells We assume the cells in the border of the domain where sensors can measure and the domain where sensors cannot measure a gateway cells. The reddish cells are gateway cells in Fig.2.

3. Supposition of the route candidate between the gateways For two different gateway cells, we assume a route between them as a route candidate. According to Assumption 1 and 2, I enumerate the route candidates for the combination of all gateway cells. Fig.3 shows the state to assume route candidates from two gateway cells. Yellowish green lines are the route candidates.

4. Calculation of the population density of each cell Using the measurement data of the laser range scanner for the constant period of time, the population density of the cells is calculated. Our method count it up how many pedestrians existed in each cell from the positional coordinate data of measured pedestrians during the measurement period (Fig.4).
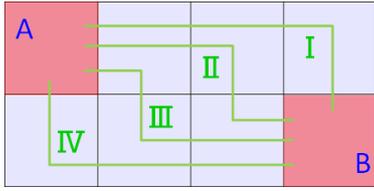
Figure 3: Assumption of Route Candidate

Because the sizes of all cells are the same, the counted number of people is considered as the population density of the cells.

Figure 4: Calculation of Density of Cells

5. Calculation of the number of sojourners   In the measurement area, some people may exist without moving (sojourners). By this method, the sojourners may be counted several times, and the population density of the cell where the sojourners are may grow large unnaturally, because sojourners are hidden behind other pedestrians. Therefore, it is necessary to exclude the data of sojourners not to influence the value of the population density. We set the minimum movement speed v_{min} for the pedestrian data succeeding for tracking, and exclude the measurement data of a pedestrian moving at a speed less than v_{min} at the time of the population density calculation.

It is necessary to know the number of sojourners because the control of sojourners is needed for the control of the pedestrian flow. Therefore our method calculate the mean number of sojourners during a measurement period based on the number of the pedestrians moving at a speed less than v_{min} which are excluded at the time of the population density calculation.

6. Estimation of the route where the pedestrian passed and its traffic   The route that the pedestrian passed is estimated from route candidates and the population density of each cell. Pedestrian is less likely to pass the cell with the small population density and pedestrian is more likely to pass the cell with the large population density. From assumption 3, when one pedestrian passed a certain route, the population of the cell which the route passes increases 1. As a result, the population density also increases 1. Based on this assumption, our method calculate the route and the traffic, so that the population density of each cell calculated from the decided quantity of flow and the decided route becomes near to the population density of each cell really measured. Because there is much number of route candidates, and the calculation

cost of verifying all combinations, we use the greedy algorithm for calculating the optimum route. The optimum route here is the route where the possibility that pedestrians pass is the highest when the population density of each cell was given. The criteria of the optimum route are as follows.

- Do not pass a cell with population density less than 0 as much as possible.

- Pass the cell which is high in population density with precedence.

The optimum route is calculated as follows by the greedy algorithm. At first, the most suitable route candidate for given population density is chosen as the temporary optimum route. The population density of each cell except the pedestrians via the temporary optimum route is calculated again. The optimum route is calculated by repeating the same process for the population density that updated a value. When one route corresponds to one pedestrian, the processing to remove pedestrians via the route corresponds to the process of reducing the population density of the cell which a route decided from Assumption 3 passes. When all the routes are decided and the population density of all cells became 0, the combination of I might calculate for the original population density might be calculated. Therefore the calculation by the greedy algorithm is repeated and finished when the total absolute value of the population density of all cells is minimized. The routes decided by the end of the processing expresses the course where each pedestrian passed. The number of each decided route expresses the number of pedestrians via the route. The route calculation processing is explained below by a specific example.

1. Calculate the most suitable course candidate as input by population density
   For example, the population density such as Fig.4 is assumed input. The number of figures expresses the population density of the cell. The route candidates in this domain are I - IV of Fig.3, and with the case of Fig.4, route IV becomes the most suitable route candidate.

2. Check the total absolute value of the population density of all cells whether the most suitable route candidate may be chosen as a route
   The current total absolute value of the population density is 40, and the total value when the route candidate IV is chosen as a route is 35. Because the total absolute value of the population density becomes small, this route candidate is chosen as a route.

3. Output the value that updated the population density of the cell which the route passes
   Because one route corresponds to one pedestrian, the population density of the cell which the chosen route passes is decreased 1. Route candidate IV is chosen in Fig.3 as a route and Fig.5 shows the value of the each cell is updated .

4. Return to (1) by the output in (3) as new input
   The whole processing mentioned above is repeated by population density of Fig.3 as new input.
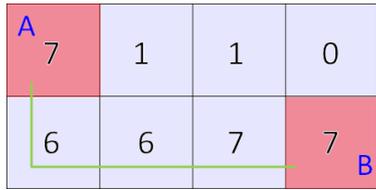
Figure 5: Population Density when route IV is Chosen
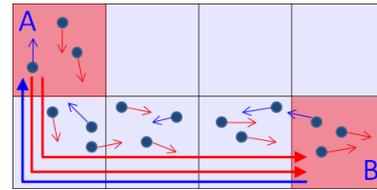
Figure 6: Determination of Direction of the Pedestrians' Movement

7. Determination of direction of the flow   The information about quantity of flow and direction are necessary for pedestrian flow model. Therefore the direction of the pedestrian's movement via the route estimated from population density is needed. Because it is difficult to estimate the movement route of the pedestrian only from the positional coordinate data (x, y), we use the tracking function of the laser range scanners. The direction of the pedestrian's movement can be estimated from the position difference of an equivalence person provided by tracking while tracking succeeds. If a certain route p between gateway cells $g1$ and $g2$ are given, the direction of the pedestrian's movement in each cell which the route passes is examined. About a pedestrian moving along the route, the number of people according to a movement direction (someone move towards $g2$ from $g1$ and someone move towards $g1$ from $g2$) is counted. From the result of counting, the direction ratio of the pedestrian's movement is calculated, and the quantity of flow of the route is divided to accord with this ratio. In this way, the quantity of flow from $g1$ to $g2$ and the quantity of flow from $g2$ to 1 on the route p are obtained.

Fig.6 shows a specific example. The points in a figure express a pedestrian, and the arrows connected to each point express the direction of the pedestrian's movement. The long arrows between gateway cell A and gateway cell B expresses routs between A and B. The quantity of flow of the route is expressed by the number of the arrow. In the example of Fig.6, the quantity of flow of the route between A and B is 3, and there are 3 walkers in each cell which the route passes. When the number of pedestrians is counted according to the direction of movement, the number of pedestrians toward cell B from cell A (the pedestrian with red arrow) is 10 and the number of pedestrians toward cell A from cell B (the pedestrian with blue arrow) is 5, because each pedestrian moves along the route. As for the direction ratio, the pedestrian toward B from A becomes 66.7% and the pedestrian toward A from B becomes 33.3%. Our method divide quantity of flow according to this ratio, and obtain a flow with 1 quantity toward A from B (blue arrow) and flows with 2 quantity toward B from A (red arrows).

The pedestrian flow model is generated in the flow form from the direction and quantity information of the flow by the above-mentioned processing.

From the measurement data which the experiment in "Whity Umeda" provided, the pedestrian flow models are as follows by the proposed method. The pedestrian flow model is generated from 10-minute measurement data until 8:26:08 a.m. on December 24, 2010 from 8:16:09 a.m. on December 24, 2010. The measurement area is a rectangular domain of $31.4m \times 27.7m$, and the size of the cell is $6.3m$ square. This area is

divided in unit cells. The number of pedestrians who existed in each unit cell for 10 minutes is counted and this number is considered as the population density of pedestrians in each unit cell. To calculate an average flow per time, the quantity of the pedestrian flow is calculated from the population density that averaged the data for 10 minutes in the area for 1 minute. Fig.7 shows the population density of each cell obtained from measurement data.
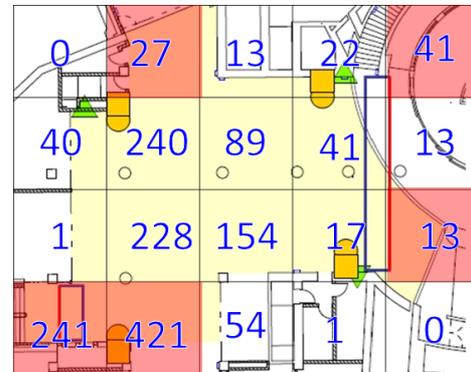
Figure 7: Population Density (8:16:9 - 8:26:8)

Fig.8 and Fig.9 show the pedestrian flow model generated from population density of Fig.7. Fig.8 is a figure expressing each pedestrian flow from a certain gateway cell to a different gateway cell. Fig.9 expresses pedestrian flow of throughout the measurement area by expressing quantity of movement to the cell which is adjacent for each cell as arrows. The direction of the arrow is the direction of the pedestrian flow, and the quantity of the pedestrian flow is proportional to the thickness of the arrow. The numbers in figures expresses the number of the sojourners of the same cell for 1 minute. The movement speed v_{min} which is standard to judge a sojourner or a pedestrian is assumed $20cm/s$.

Furthermore, another pedestrian flow model is generated from the measurement data until 6:36:08 p.m. on December 24, 2010 from 6:26:09 p.m. on December 24, 2010 on the equal conditions. Fig.10 and Fig.11 show the generated pedestrian flow model.

Because there are many people moving for commuting and attending school in the time of the morning, the characteristic flow that there is much movement quantity for the same direction is seen. People toward various directions are detected in the evening time. In addition, the existence of the sojourner is outstanding because there are many people to stop in open spaces such as the measurement area.
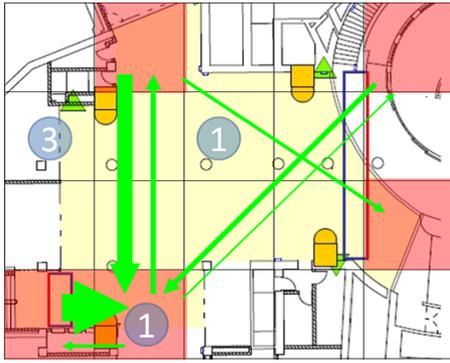
Figure 8: Pedestrian Flow (Gateway to Gateway)
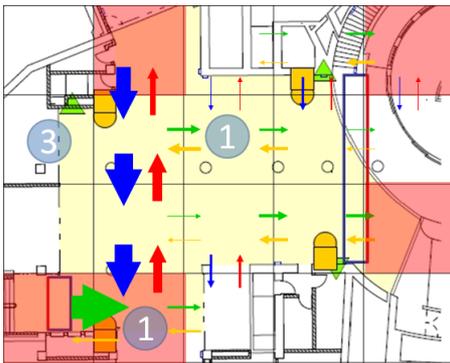
Figure 9: Pedestrian Flow (Cell to Cell)
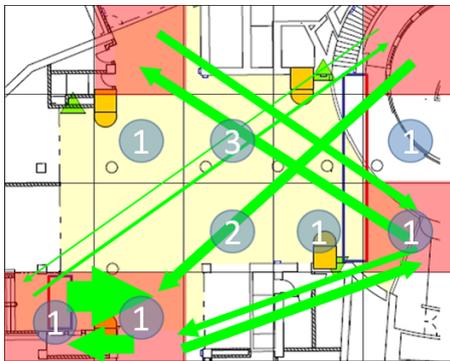
Figure 10: Pedestrian Flow (Gateway to Gateway)
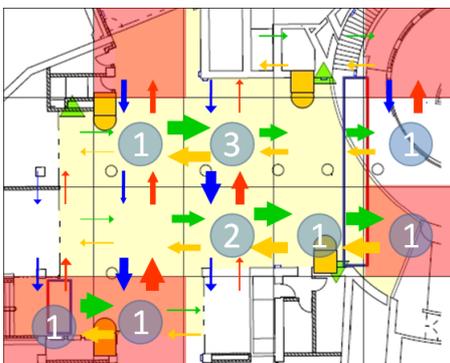
Figure 11: Pedestrian Flow (Cell to Cell)

We also reproduced the people moved along flows based on the flow on Fig.9 and Fig.11, using MobiREAL Animator[13] realistically to output comprehensible animation data visually.

5 PERFORMANCE EVALUATION

Because of the failure in the tracking of the pedestrian by occlusion, the measured data which completely caught all pedestrians in the measurement area cannot be acquired using laser range scanners. Therefore, we artificially generate the scenario data which the behavior of all pedestrians can completely reproduced based on actual survey data. And the pedestrian flow model is generated using the proposed method for this scenario data. By comparing this produced pedestrian flow model with the scenario data, we evaluate the reproducibility of the proposed pedestrian flow model to the scenario data. We compared the flow quantity toward exit cell g_2 from entrance cell g_1, and examined how much quantity of flow could reproduced.

The scenario data is generated from the data measured in "Whity Umeda". The position of gateway cells and the size of each cell make setting same as Fig. 2. Fig.12 expresses flows of the movement of pedestrian of the scenario data which made from the measured data of about 8:00 a.m. of December 24, 2010. The reddish cells in the figure are gateway cell (A - E), and the black arrow expresses the movement of pedestrians of scenario data. The thickness of the arrow expresses to the flow quantity of pedestrians. There are particularly many pedestrians moving between A and D or C and D in this scenario. We evaluate the precision of the proposed model for the state with many pedestrians who move to top and bottom right and left for a cell. Fig.13 shows the comparison between scenario data of Fig.12 and the flow quantity of routes linking each gateway cell of the generated pedestrian flow model. Some false detection occurs, but the average reproduction rate of the pedestrian flow quantity is about 82.9%, and pedestrian flow can be reproduced almost definitely by the proposed method.
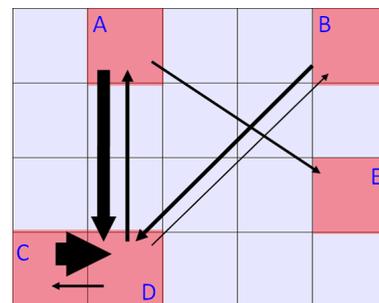
Figure 12: Scenario Modeling from Measured Data in the Morning

Next, we evaluate the model using the scenario data which made from the measured data of about 6:00 p.m. of December 24, 2010. Fig.14 expresses flows of the movement of pedestrian of the scenario data. There are particularly many
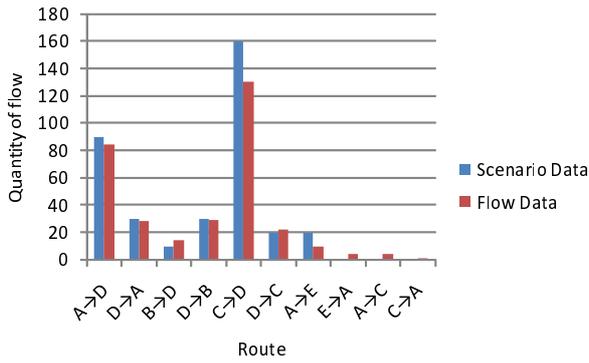
Figure 13: Comparing Scenario Data and Generated Flow

pedestrians moving between A and E or B and D in this scenario. We evaluate the precision of the proposed model for the state with many pedestrians who move for a cell diagonally. Fig.15 shows the comparison between scenario data of Fig.14 and the flow quantity of routes linking each gateway cell of the generated pedestrian flow model. The average reproduction rate of the pedestrian flow quantity is about 82.9%, and pedestrian flow can be also reproduced almost definitely in this case.
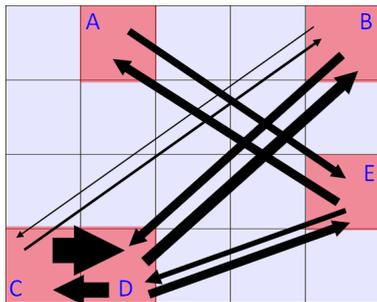
Figure 14: Scenario Modeling from Measured Data in the Evening
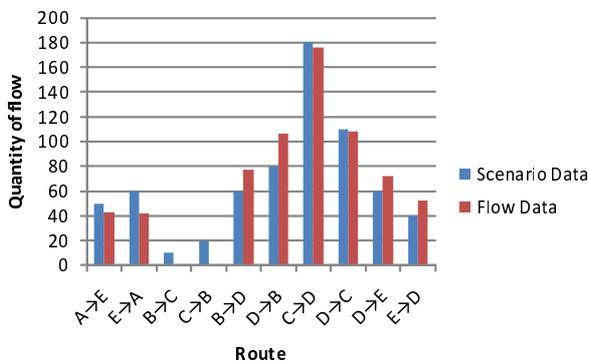
Figure 15: Comparing Scenario Data and Generated Flow

Finally, we evaluate the model using the scenario data which made from the measured data of about 1:00 p.m. of December

24, 2010. Fig.16 expresses flows of the movement of pedestrian of the scenario data. Pedestrians move to various courses in this scenario. We evaluate the precision of the proposed model for the state that there are the moving pedestrians to various courses at the same time. Fig.17 shows the comparison between scenario data of Fig.16 and the flow quantity of routes linking each gateway cell of the generated pedestrian flow model. The average reproduction rate of the pedestrian flow quantity is about 59.6%. In this case, pedestrian flow cannot be reproduced definitely. This depends on that it is difficult to distinguish a pedestrian moving between C and E, B and C from a pedestrian moving between D and E, B and D respectively, because gate way cell C and D are adjacent.
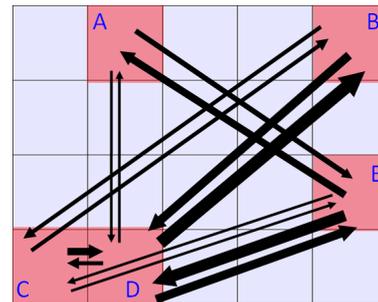
Figure 16: Scenario Modeling from Measured Data in the Afternoon
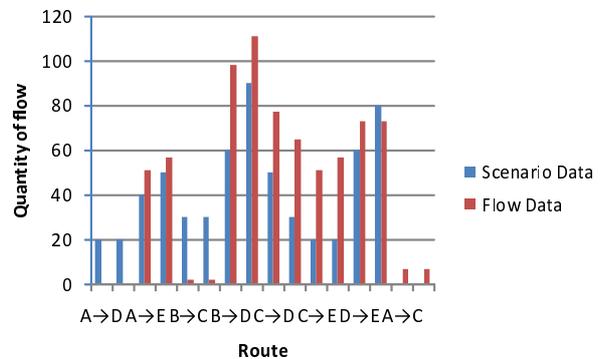
Figure 17: Comparing Scenario Data and Generated Flow

6 CONCLUSION

In this paper, we proposed the method for modeling of pedestrian flow in the space consisting of some gateways, some passages and various obstacles such as underground shopping center. The proposed method estimates the routes that the pedestrians would have passed not from the individual positional data of each pedestrian but from the population density at each cell of the measurement area. As a result, the proposed method can generate the pedestrian flow model from the data that the all behaviors of all pedestrians are not observed.

Future work include the model generation using the sojourner information of pedestrians, and the revision of the population density of each cell by the complement of the pedestrians who cannot observe by laser range scanner.

REFERENCES

[1] Josh Broch, David A.Maltz, David B. Johnson, Yih-Chun Hu and Jorjeta Jetcheva, "A Performance Comparison of Multi-Hop Wireless Ad Hoc Network Routing Protocols," in Proceedings of the 4th Annual ACM/IEEE International Conference on Mobile Computing and Networking(MobiCom'98), pp.85–97 (1998).

[2] Esa Hyytia, Pasi Lassila and Jorma Virtamo, "Spatial Node Distribution of the Random Waypoint Mobility Model with Applications," IEEE Transactions on Mobile Computing, Vol.5, No.6, pp.680–694 (2006).

[3] Christian Bettstetter, "Mobility modeling in wireless networks: Categorization, smooth movement, and border effects," ACM SIGMOBILE Mobile Computing and Communications Review, Vol.5, No.3, pp.55–67 (2001).

[4] Tommy Chu and Ioanis Nikolaidis, "Node density and connectivity properties of the random waypoint model," Computer Communications, Vol.27, No.10, pp.914–922 (2004).

[5] Andres Rojas, Philip Branch and Grenville Armitag, "Experimental validation of the random waypoint mobility model through a real world mobility trace for large geographical areas," in Proceedings of the 8th ACM/IEEE International Symposium on Modeling, Analysis and Simulation of Wireless and Mobile Systems(MSWiM2005), pp.174–177 (2005).

[6] Amit Jardosh, Elizabeth M. Belding Royer, Kevin C. Almeroth and Subhash Suri, "Towards Realistic Mobility Models For Mobile Ad hoc Networks," in Proceedings of the 9th Annual ACM/IEEE International Conference on Mobile Computing and Networking(MobiCom2003), pp.217–229 (2003).

[7] Matthias Hollick, Tronje Krop, Jens Schmitt, Hans-Peter Huth and Ralf Steinmetz, "Modeling Mobility and Workload for Wireless Metropolitan Area Networks," Computer Communications, Vol.27, No.8, pp.751–761 (2004).

[8] Wei-Jen Hsu, Kashyap Merchant, Haw-Wei Shu, Chih-Hsin Hsu and Ahmed Helmy, "Weighted waypoint mobility model and its impact on ad hoc networks," ACM SIGMOBILE Mobile Computing and Communications Review, Vol.9, No.1, pp.59–63 (2005).

[9] Minkyong Kim, David Kotz and Songkuk Kim, "Extracting a mobility model from real user traces," in Proceedings of the 25th IEEE International Conference on Computer Communications(Infocom2006), pp.1–13 (2006).

[10] Kumiko Maeda, Akira Uchiyama, Takaaki Umedu, Hirozumi Yamaguchi, Keiichi Yasumoto and Teruo Higashino, "Urban Pedestrian Mobility for Mobile Wireless Network Simulation," Ad Hoc Networks, Elsevier, Vol.7, No.1, pp.153–170 (2009).

[11] Huijing Zhao and Ryosuke Shibasaki, "A Novel System for Tracking Pedestrians Using Multiple Single-Row Laser-Range Scanners," IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans, Vol.35 No.2, pp.283–291 (2005).

[12] UTM-30LX, `http://www.hokuyo-aut.co.jp/02sensor/07scanner/utm_30lx.html`.

[13] MobiREAL, `http://www.mobireal.net/`.

[14] Ajo Fod, Andrew Howard and Maja J.Mataric, "Laser-Based People Tracking," in Proceedings of the 2002 IEEE International Conference on Robotics and Automation(ICRA2002), pp.3024–3W029 (2002).

[15] Dominik Joho, Christian Plagemann and Wolfram Burgard, "Modeling RFID signal strength and tag detection for localization and mapping," in Proceedings of the 2009 IEEE International Conference on Robotics and Automation(ICRA2009), pp.3160–3165 (2009).

[16] Jae Lee, Takashi Tsubouchi, Kenjiro Yamamoto and Saku Egawa, "People Tracking Using a Robot in Motion with Laser Range Finder," in Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems2006(IROS2006), pp.2936–2942 (2006).

[17] Hirohiko Kawata, Akihisa Ohya, Shin'ichi Yuta, Wagle Santosh and Toshihiro Mori, "Development of ultra-small lightweight optical range sensor system," in Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems 2005(IROS2005), pp.1078-1083 (2005).

[18] Ohba Ryohji, Yamashita Tomohisa, Ukai Osamu and Kato Shinsuke, "Development of Hazard Prediction System for Intentional Attacks in Urban Areas," in Proceedings of the 7th International Symposium on New Technologies for Urban Safety of Mega Cities in Asia(USMCA2008), pp.687–696 (2008).

Estimation of Presence Information of a User

by using RSSI of Bluetooth and Acceleration Sensor Data

Yoh Shiraishi[*], Masataka Kikawa[**] and Osamu Takahashi[*]

[*] School of Systems Information Science, Future University Hakodate, Japan
[**] Graduate School of Systems Information Science, Future University Hakodate, Japan
{siraisi, g2109011, osamu}@fun.ac.jp

Abstract - In recent years, various kinds of services using presence information that shows user's activity situation and location information are increasing. We focus presence estimation about "remaining and leaving judgment" from presence states with higher value. "Remaining and leaving judgment" is to distinguish "remaining" and "leaving" a seat at the front of a PC desk. In this paper, we propose a method to estimate presence information ("remaining" or "leaving" a seat) of a user by using RSSI of a Bluetooth device and a 3-axis acceleration sensor. We conducted some experiments to evaluate the proposed method. The experimental results showed that our method can satisfy the requested requirements.

Keywords: presence estimation, context awareness, RSSI, Bluetooth, acceleration sensor.

1 INTRODUCTION

Recently, various kinds of services using presence information [1, 2, 3] that shows user's activity situation and location information are increasing. Such presence information is useful information that can add value to the services and will be applicable for a wide range of fields, because the presence information to express various kinds of conditions of the user can be updated and used in real-time. However, in the existing approaches, the presence state is changed manually, and we can see reliability degradation of the presence information caused by some human mistakes such as a failure and delay in updating presence information. Therefore, in recent years, researches about presence estimation by using context awareness technologies such as sensing technologies and wireless radio technologies have much attention.

In this study, we tackle the presence estimation problem on two points: The first point is "to estimate presence information with higher added value." We can classify the presence information to one with higher value and another with lower value based on the utility frequency and situation of the presence information, because the presence information includes various kinds of states for applying many situations. We focus presence estimation about "remaining and leaving judgment" from presence states with higher value. "Remaining and leaving judgment" is to distinguish "remaining" and "leaving" a seat at the front of a PC desk. The second point is "to improve the judgment accuracy and usability." If general consumers as well as corporation systems use the presence information

aggressively, such information can be applicable for various kinds of scenes. Therefore, we decided to take into consideration not only the improvement in accuracy of presence estimation but also practicality for considering ease to use.

Considering the above two points, in this paper, we propose a method to estimate presence information ("remaining" or "leaving" a seat) of a user by using RSSI of a Bluetooth device and a 3-axis acceleration sensor. We conducted some experiments to evaluate the proposed method. Our method uses a mobile terminal equipped with Bluetooth and a 3-axis acceleration sensor, and a desktop computer connected to the Bluetooth module.

This paper organizes as follows. In section 2, we refer to the related works and point out some problems. In section 3, we propose a new method for context awareness applications and describe our method in details. In section 4, we show the experimental results. Finally, we conclude in section 5.

2 RELATED WORK AND APPROACHES

2.1 Presence information estimation

"Remaining and leaving presence estimation" is to judge whether a user remains or leaves a seat in front of his (her) desk while operating a computer on the desk. Such presence information can be estimated by using various kinds of devices equipped with a desktop computer and a mobile terminal. There are typical sensing technologies to estimate "remaining and leaving presence": Bluetooth, a USB key device and a sensor device [4, 5].

Bluetooth is a wireless communication technology and is equipped to many devices such as mobile terminals, laptop computer and tablet PCs. The wireless technology can be easily used by consumers and is power-saving technology lower than other wireless technologies such as wireless LAN, and has the advantage of lower radio wave interference. Because Bluetooth is such good properties, there are many existing researches and products for presence estimation using Bluetooth [6]. Also, we proposed and evaluated a method for estimating remaining and leaving presence information [7]. Such methods using Bluetooth for remaining and leaving presence estimation are classified into three methods.

a) Inquiry method
This method is a method using two functions (search by a master device of Bluetooth and response by a slave device) when the connection between these Bluetooth devices is established. In this case, misjudgment about remaining presence ("remaining misjudgment") occurred because of dropping search/response packets by surrounding environmental changes such as obstacles and radio interference. Remaining misjudgment is to label the presence of a remaining user as "leaving."

b) Communication link method
This method is to monitor communication links (connection status) while a master device is connecting with slave devices. In this case, misjudgment about leaving presence ("leaving misjudgment)" occurs because of the increase of radio field strength of communication links. Leaving misjudgment is to label the presence of a leaving as "remaining."

c) RSSI method
This method uses RSSI (radio field strength) of a Bluetooth. RSSI expresses the strength of the received radio signal as a numerical value and can be used an index that expresses a relative distance between devices connected with wireless communication.

2.2 Requirements

We try to propose a presence estimation method with high accuracy and practical use. To realize such method, we have to consider the following points:

1) Device selection: we have to select devices and technologies with high penetration rate.
2) Estimation accuracy: we have to propose a method with adequate estimation accuracy that can satisfy the requested requirements.
3) Environmental constraints: we have to relax experimental constraints.

2.2.1. Device selection

The penetration rate of devices and technologies that we select is important in case that we consider the practicality of the proposed method. Using special devices with high cost can improve the accuracy of presence estimation. However, it is not practical in order to be used by many consumers. To estimate presence by using devices and technologies with high penetration rate is desirable.

2.2.2. Estimation accuracy

In order to estimate presence information, a method with high accuracy that can judge the state of a user correctly is required. We define the judgment of remaining/leaving presence with high accuracy as that both accuracy of "remaining judgment" and "leaving judgment" are high. Concretely, we aim to satisfy the following requirements with over 95 percent.

● Cond-1: Condition about remaining judgment
In case that a user is working in front of or near a PC desk, an estimation method has to judge the user state as "remaining." We define "remaining area" as an area for remaining judgment. The relative distance between a user in remaining area and PC is from 0 to 1 meter (as shown in Figure 1). If a system judges as "leavening" although a user is in remaining area, the judgment is wrong.

● Cond-2: condition about leaving judgment
After a user starts leaving a seat, when the user reaches at adequate distance, the system has to judge the user state as "leaving." We define "leaving judgment area" as an area for leaving judgment. The relative distance between a user in leaving judgment area and PC is from 1 to 10 meters (Figure 1). Also, when the distance between the user and PC is over 10 meters, the system necessarily judges the user state as "leaving". If the system judges as "remaining" when the user is in the leaving area, the judgment is wrong.
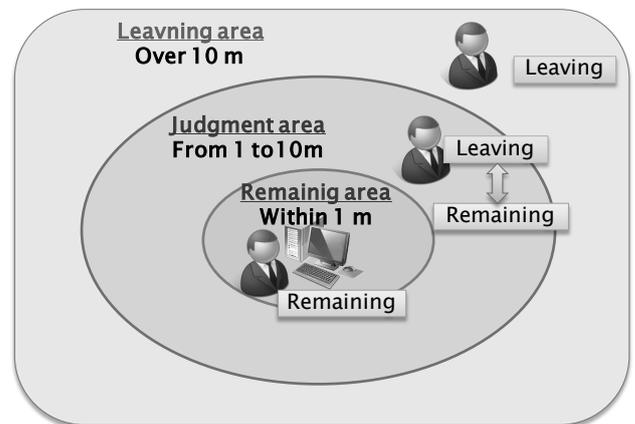
Figure 1: Area requirements

These conditions are more difficult than those of the existing method that we proposed in [7]. In the existing work, the distance between PC and a user in remaining area is 0 to 10 meters, and the distance between PC and a user in leaving area is over 30 meters. The requirements of this paper are more practical.

2.2.3. Environmental constraints

We proposed a method for remaining and leaving presence estimation based RSSI method [7]. This method solved the misjudgment problem of RSSI method under two experimental constraints: "to fix a Bluetooth dongle with a desktop computer on a desk" and "to fix position of a Bluetooth device with a user". There is the possibility that the usability of the method decreases by fixing the position and type of Bluetooth devices. To use the estimation method without regard to such constraints is desirable.

2.3 Approach

There are the following problems based on RSSI of a Bluetooth:

1) Variation in leaving distance.
   RSSIs in different environments decrease in different manners. Consequently, if we estimate remaining / leaving presence by comparing RSSI with the specified threshold, variation in the leaving distance when the presence is judged as "leaving" occurs.
2) Misjudgment in remaining area.
   In case that the threshold is set at high-level in order to keep the leaving distance constant, the misjudgment that the presence is judged as "leaving" in remaining area occurs.

Our existing study [6] introduced some constraints about positions and type of Bluetooth devices to solve these problems.

In order to solve these problems without setting a limitation about Bluetooth devices, we introduce different viewpoints different from the existing researches.

a) Waling judgment in remaining area
   It is expected that misjudgment in remaining area can be reduced by distinguishing between "walking" and "stopping" of a user by using technologies except Bluetooth. Therefore, we adopt a 3-axis acceleration sensor that is widely used for action estimation. Also, a 3-axis acceleration sensor has a high rate of mount on mobile terminals similar to a Bluetooth device.
b) Comparison of RSSI and variable threshold at leaving.
   We examined an introduction of the variable threshold not the fixed threshold to deal with a problem about variation in the leaving distance. We change the threshold in real-time depending on every user environments, in order to keep the leaving distance in constant. Also, the variable threshold is set when the state of a user is changed from "stopping" to "walking".

3 PROPOSED METHOD

3.1 Overview of the proposed system

Figure 2 shows an overview of the proposed system.
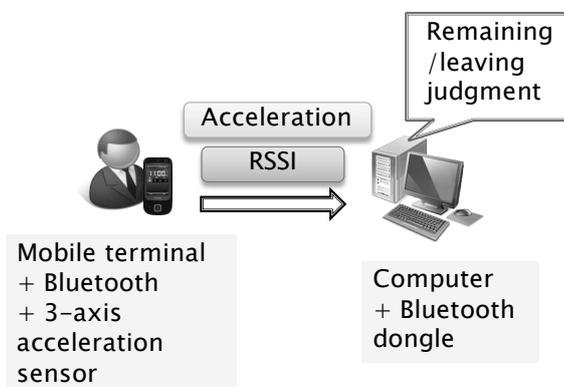
Figure 2: An overview of the proposed system

Our system consists of a mobile terminal with Bluetooth and a 3-axis acceleration sensor and a computer connected with Bluetooth dongle. A process for remaining and leaving judgment runs on the computer that receives RSSI and acceleration sensor data from a mobile terminal through Bluetooth.

The proposed method executes "attending and leaving judgment" with two phases: "walking judgment" (Step 1) and "relative distance judgment" (Step 2).

In "walking judgment" phase (Step 1), a walk judgment of a user is performed by extracting the amount of the features from acceleration sensor data through FFT (Fast Fourier Transform) as the process during "remaining a seat". If it detects that a user is walking, it moves to "relative distance judgment" phase (Step 2). At Step 2, a user's condition ("remaining" or "leaving" a seat) is discriminated by measuring the relative distance between the desktop computer and a user from the amount of the features of RSSI as the process during "leaving a seat."

3.2 Walking judgment by acceleration sensor

In "walking judgment" phase, our system receives 3-axis acceleration sensor data from a mobile terminal, calculates the feature quantities from these acceleration data by using FFT. An acceleration sensor is often used for action estimation in the field of ubiquitous and mobile computing [8-10]. It notes that we can see a peak near 2 Hz in FFT results for "walking" action. Also, the power spectrum is low for stopping. Therefore, we compare the power spectrum near 2 Hz with the pre-specified threshold for walking judgment. In the proposed method, the sampling rate of FFT is 40 Hz and the feature quantities are calculated by using latest data for 3 seconds (128 samples).

The walking judgment phase is executed as follows:
[A-1] This phase is transited from "remaining" state.
[A-2] A system judges whether a user is walking or not by examining acceleration sensor data from a mobile terminal.
[A-3] If the state of the user is "walking", the user starts moving away from his or her PC desk. Go to the relative distance judgment phase. Otherwise, go to [A-4].
[A-4] If the state of the user is "stopping", the user does not move from his or her desk and the state is judged as "walking". Go to [A-2].

3.3 Relative distance judgment by RSSI of Bluetooth

Our method sets the variable threshold based on latest RSSI data immediately after switching from "stopping" to "walking" in the walking judgment phase. By doing so, it calculates the optimal threshold depending on the user environment when starting the "leaving" action. The environments have different properties such as the position and holding state of a mobile terminal, materials of a desk and the type of a device.

In the proposed method, we define the average value of this latest RSSI data for 1 second as "baseline." The baseline expresses the RSSI value of the moment a user starts walking and changes depending on the differences of the environments. The proposed method sets the line decreased from the baseline as "variable threshold." The degree of the decrease is given as the "RSSI threshold level" that is suitable for the requirements.

The relative distance judgment phase is executed as follows:

[B-1] This phase is transited from the walking phase judgment.

[B-2] The system calculates and sets the baseline and the variable threshold.

[B-3] It measures the relative distance by using RSSI of Bluetooth.

[B-4] If the latest RSSI value is lower than the variable value, the state of the user is judged as "remaining". Go to [B-3] (the measurement process for the relative distance is repeated). Otherwise go to [B-5]

[B-5] If the latest RSSI value is between the variable threshold and the baseline, go to [B-3]. The measurement of the relative distance is repeated with "remaining" state. Otherwise go to [B-6]

[B-6] If the latest RSSI value is higher than the baseline, change from "leaving" state to "remaining" state, and go to the walking judgment phase.

4 EXPERIMENTS AND RESULTS

4.1 Experimental settings

4.1.1. Data acquisition system for evaluation

We constructed a data acquisition system for evaluation. The system consists of a mobile terminal and a computer shown in Figure 3. We used HT-03A (HTC) as a mobile terminal and equipped 3-axis acceleration sensor WAA-001 (wireless technology, inc.) to the back face of the terminal. WAA-001 is an acceleration device capable for Bluetooth communication. A Bluetooth dongle is connected to a computer. The computer with the dongle can communicates with the mobile terminal. The data acquisition system on the computer was implemented by using C# language and can acquire Bluetooth's RSSI and acceleration sensor data from the terminal.
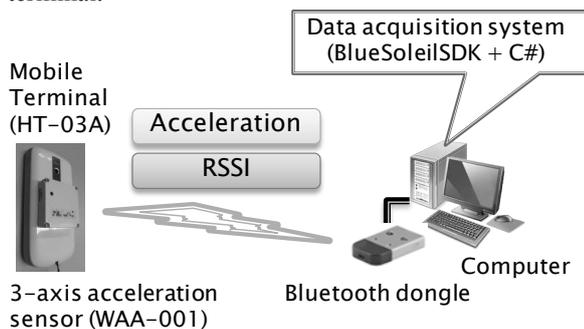
Figure 3: Data acquisition system

4.1.2. Experimental scenario

In this experiment for evaluation, an examinee behaves along an experimental scenario that consists of the following actions:

(Action 1) Remaining a seat
(Action 2) Leaving a seat
(Action 3) Returning a seat

First, an examinee remains a seat in front of a PC desk for 75 seconds (Action 1). During Action 1, the examinee changes the state of a mobile terminal (how to hold the terminal) at every 15 seconds. The state of a mobile terminal in remaining a seat is the following 5 patterns:

- On a desk
- In a breast pocket of a shirt
- In a front pocket of pants
- Main operation
- Calling

Next, the examinee leaves from the PC desk up to 10 meters (Action 2). The examinee moves at 1m/s for 10 seconds. The state of the mobile terminal in leaving the seat is the following 4 patterns:

- In a breast pocket of a shirt
- In a front pocket of pants
- Main operation
- Calling

Finally, the examinee returns to the seat along the same route as Action 2 (Action 3). The length of the route is 10 meters. During Action 3, the examinee moves at 1 m/s for 10 seconds.

A sequence of Action 1-3 is a trial of this experiment. By using our developed data acquisition system, we collected Bluetooth's RSSI data and 3-axis acceleration sensor data for a total of 95 seconds.

4.1.3. User environments for this experiment

We prepared 4 user environments with different characteristics: UE-1 (Figure 4), UE-2 (Figure 5), UE-3 (Figure 6) and UE-4 (Figure 7).

Figure 4: User environment 1 (UE-1)

Figure 5: User environment 2 (UE-2)
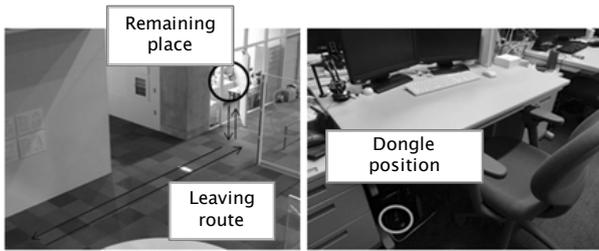
Figure 6: User environment 3 (UE-3)
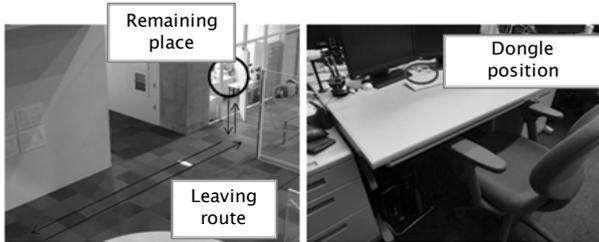
Figure 7: User environment 4 (UE-4)

Table 1 shows the characteristics of each environment. The remaining place of the user, the route in leaving of the user, and the position of the Bluetooth dongle are shown in each figure.

Table 1: User environments with different characteristics

	Materials of desk	computer
UE-1	Iron desk	Laptop
UE-2	Wooden desk	Laptop
UE-3	Wooden desk, Iron drawer	Desktop
UE-4	Wooden desk, Iron drawer	Desktop

Also, we prepared two kinds of Bluetooth dongles by PLANEX COMMUNICATIONS (as shown in Table 2). These Bluetooth devices belong to different Bluetooth Class.

Table 2: Specifications of Bluetooth dongles

Product name	BT-MicroEDRIX	BT-MicroEDR2X
Bluetooth Class	Class 1	Class 2
Max. comm. distance	100 meter (theoretical)	10 meter (theoretical)
Transmission power	About 100 mW	About 2 mW
Max. comm. speed	2.1 Mbps	2.1 Mbps

In this experiment, we have 4 user environments, 2 kinds of Bluetooth dongles, and 4 patterns in how to hold a mobile phone (there are no more than 4 patterns in Action 2 and 3, although there are 5 patterns in Action 1). A experimental trial under a combination of these parameters is repeated five times. Consequently, we had 160 trials and collected the experimental data through our data acquisition system.

4.2 Experimental results

We examined the accuracy about remaining and leaving presence by judgment using data collected through the data acquisition system. This experiment has two parameters:

The threshold values for walking judgment are 8, 9, 10, 12 W/Hz. The RSSI threshold levels are -10, -9, -5, -4, -3 dB.

Table 3 shows the accuracy of judgment of remaining presence. The accuracy is the rate of the correct judgments for all trials.

Table 3: Accuracy of remaining judgment (unit: percent)

RSSI threshold level (dB)	Threshold for walking judgment (W/Hz)			
	8	9	10	12
-3	97.5	98.8	99.4	100
-4	98.8	99.4	99.4	100
-5	98.8	99.4	99.4	100
-8	98.8	100	100	100
-10	98.8	100	100	100

Table 4 shows the accuracy of judgment of remaining presence.

Table 4: Accuracy of leaving judgment (unit: percent)

RSSI threshold level (dB)	Threshold for walking judgment (W/Hz)			
	8	9	10	12
-3	100	100	97.5	96.3
-4	100	99.4	96.9	96.3
-5	100	99.4	96.9	96.3
-8	94.4	95.6	94.4	93.8
-10	91.9	92.5	93.1	90.6

The experimental results (Table 3 and 4) show that our method can satisfy the requested conditions (Cond-1 and Cond-2). The accuracy of remaining and leaving judgment is over 95 percent in almost combination of the threshold for walking judgment and RSSI threshold level. Table 5 shows examples of the optimal values with high accuracy. The accuracy of these combinations is over 98.8 percent in remaining judgment as well as leaving judgment.

Table 5: Examples of optimal values

Threshold for walking judgment (W/Hz)	RSSI threshold level (dB)
8	-4
8	-5
9	-3
9	-4
9	-5

5 CONCLUSION

In this paper, we examined a presence-estimation method that can add value to various kinds of ubiquitous services. Especially, we pay attention to remaining and leaving presence information and proposed a method for judging whether a user is remaining and leaving his or her seat in front of user's desk. Our system consists of a mobile terminal equipped with a 3-axis acceleration sensor and Bluetooth device and a computer with Bluetooth dongle. To solve misjudgment problems in existing research, our

method executes two kinds of process: "walking" judgment and the relative distance judgment.

We conducted evaluation experiments in various kinds of environments/conditions in order to verify the usefulness of the proposed method. The experimental results showed that the proposed method could perform "remaining and leaving judgment" with high accuracy under various conditions. The accuracy of "remaining" judgment was 98 percent or more, and the accuracy of "leaving judgment" was 100 percent. Moreover, we analyzed the distance of "leaving judgment". As a result of our analysis, we confirmed that the proposed method executed the leaving process within 5 meters in most cases, although our initial goal is to execute the leaving process within 10 meters. This result means that our method achieved the accuracy largely exceeding the initial goal. Since our method has no constraints when using it, we think that this method is one with high usability and the Usefulness of the method is high when assuming the real environments. For the future, to improve estimate accuracy of sensor data, we consider a database system dealing participatory sensing data. And to realize this system, we consider the method to deal mobile sensing data.

REFERENCES

[1]   J. Rosenberg and H. Sugano, "A Model for Presence and Instant Messaging", http://www.ietf.org/rfc/rfc2778.txt, 2000.

[2]   H. Sugano, S. Fujimoto, G. Klyne, A. Bateman, W. Carr and J. Peterson, "Presence Information Data Format (PIDF)," http://www.ietf.org/rfc/rfc3863.txt, 2004.

[3]   H. Hchulzrinne, V. Gurbani, P. Kyzivat, J. Rosenberg, "RPID : Rich Presence Extensions to the Presence Information Data Format (PIDF)," http://www.ietf.org/rfc/rfc4480.txt, 2006.

[4]   BlueSoleil，"BlueLock"，BlueSoleil Download，http://www.bluesoleil.com/，2008/07.

[5]   IO DATA, "WinSensor," WinSensor series, http://www.iodata.jp/, 2011/01.

[6]   Nanao, "FlexScan LCD," FlexScan LCD, http://www.eizo.co.jp/, 2011/01.

[7]   M. Kikawa, T. Yoshikawa, S. Ookubo, A. Takeshita, Yoh Shiraishi and Osamu Takahashi, "A Presence-detection Method using RSSI of a Bluetooth Device," International Journal of Informatic Society (IJIS), Vol.2, No.1, pp.23-31, 2010.

[8]   N. Kern, S. Antifakos, B. Schiele and A. Schwaninger, "A Model for Human Interruptability: Experimental Evaluation and Automatic Estimation from Wearable Sensors," Proc. of 8th IEEE International Symposium on Wearable Computers（ISQC '04）, pp.158-165, 2005.

[9]   N. Ravi, P. Mysore, Michael L. Littman, "Activity Recognition from Accelerometer Data," Proc. of 7th Conference on Innovation Application of Article Intelligence, pp.1541-1546, 2005.

[10] L. Bao and S. S. Intelle, "Activity Recognition from User-Annotated Acceleration Data," PERVASIVE 2004, pp.1-17, 2004.

Keynote Speech:
Prof. Peter T. Kirstein
(University College London, UK)

IPv6 status and Uses

Peter T. Kirstein

University College London

1

19 June 2011 IWIN Workshop

The Theme of the Talk

- **There are many current considerations of the shape of the future Internet**
  - **The only immediate candidate is IPv6**
- **Information on what it is, why it will come, and what it will bring are vital to all**
  - **But the planning for the transition is slow**
- **Many regions are ready to move in this direction**
  - **I will indicate what is happening here**
- **The transition will require training**
  - **Some of the training initiatives will be mentioned here**
- **What are the advantages, and why must it happen?**

2

19 June 2011 IWIN Workshop

Contents of Talk

- **The Internet and IPv4 and address depletion**
- **What is IPv6, and its advantages**
- **How near is transition to IPv6?**
- **New Protocols and IPv6 drivers**
- **Some important applications that would benefit**
- **Training and 6DEPLOY**

3

19 June 2011 IWIN Workshop

Contents of Talk

- ***The Internet and IPv4 and address depletion***
- **What is IPv6, and its advantages**
- **How near is transition to IPv6?**
- **New Protocols and IPv6 drivers**
- **Some important applications that would benefit**
- **Training and 6DEPLOY**

4

19 June 2011 IWIN Workshop

IPv4, IPv6 and Address Depletion

- **Because of its 32 bit address length, IPv4 has 256 blocks of 16 M addresses – called a /8**
- **There is a mechanism for allocating the addresses**
  - **But they are being used up at a rapid rate**
- **There are many aspects of IPv4 protocols that one thinks now need improvement**
  - **But address depletion is an important driver**
- **We decided around 1990 that a re-think of current IPv4 protocols was needed**
  - **Result was IPv6**
  - **Much larger address space and other improvements**

19 June 2011   IWIN Workshop

5

Mechanism of IPv4 Address Allocation

- **IPv4 addresses used to be allocated in an *ad hoc* fashion**
  - **I personally held two /8 blocks for .UK and .Int!**
- **Now there is a system of Internet Registries**
  - **World (IANA), Regional (RIRs), and Local (LIRs)**
  - **RIRs allocate blocks to Local Internet Registries (LIRs)**
  - **LIRs allocate blocks to end users**
- **IANA allocates /8 blocks to RIRs – No charge**
  - **RIRs have their own policies on such allocations**
  - **LIRs have their own policies subject to some RIR rules**
  - **All provisions are only on a cost recovery basis**
- **See http://www.nro.net/documents/comp-pol.html**

19 June 2011   IWIN Workshop

6

Allocations of /8 Addresses 3/09
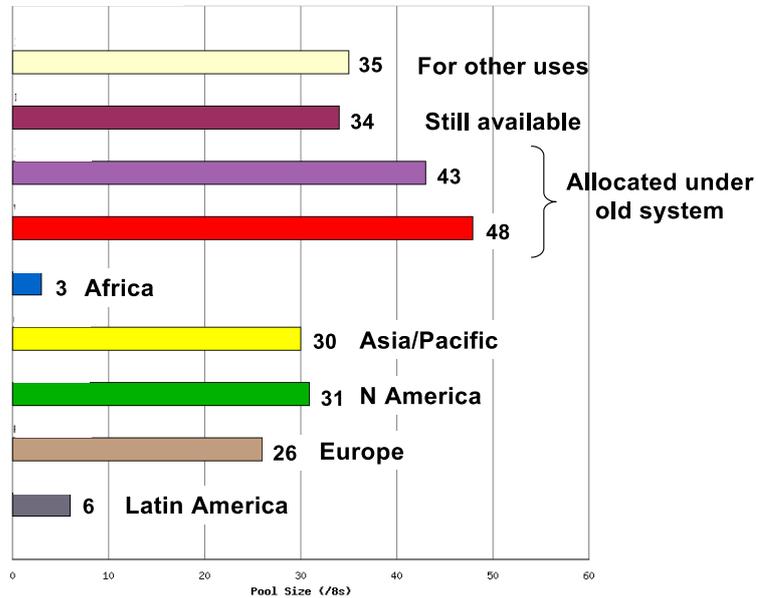
7

IPv4 Allocations by Region
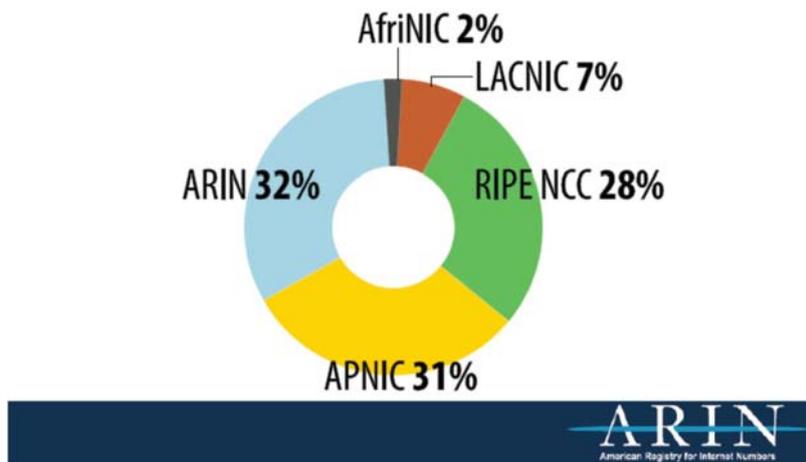
8

Estimate of RIR Exhaustion Date

- **Exhaustion defined as when last /8 is reached**
- **When exhaustion date reached much firmer allocation policies are introduced**
- **Estimated Exhaustion Dates (at current rates)**
  - **APNIC:19-Apr-2011**
  - **RIPENCC:28-Feb-2012**
  - **AFRINIC:31-Jul-2013**
  - **ARIN:20-Nov-2013**
  - **LACNIC:08-May-2014**
- **Exhaustion date does not mean end, but much stricter rules on allocation – e.g. at most 1 K address**

9

19 June 2011   IWIN Workshop

IPv4 Allocations by Region – 2/11

10
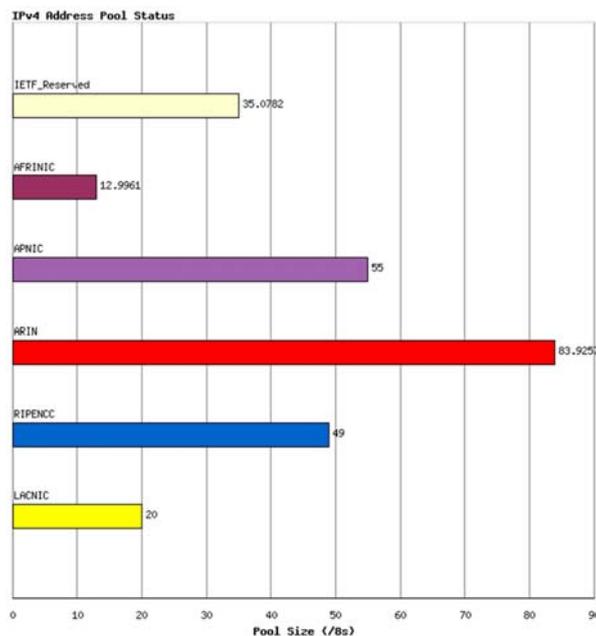
19 June 2011   IWIN Workshop

Exhaustion Date Estimate Variance
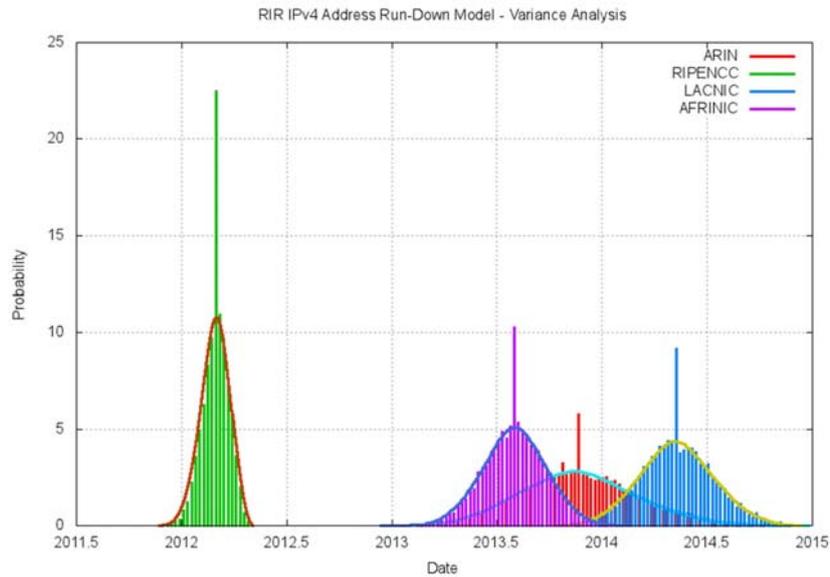
IPv4 /8s In Various Categories
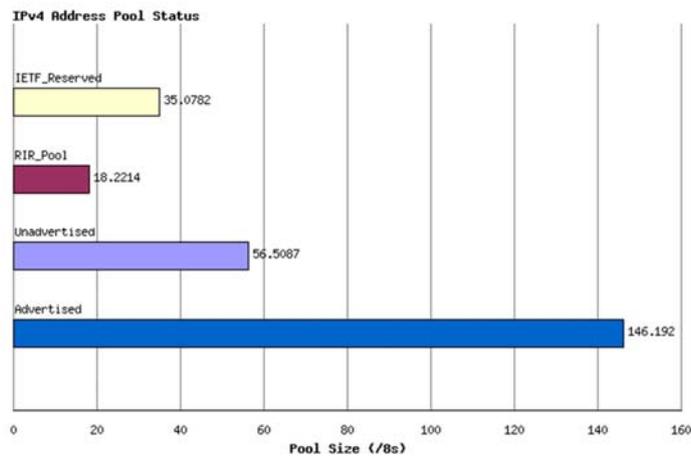
Urgency due to Address Depletion

- **Address depletion stated too often**
  - **but now there**
  - **http://www.potaroo.net/tools/ipv4/**
- **Address Exhaustion when lLast /8 reached**
- **Triggers new emergency mechanisms to conserve space**
- **IANA reached last 5 /8s in February 11**
  - **Allocated last 5 /6 blocks to RIRs**
  - **APNIC reached last /8 in April 2011**
- **Will only get much smaller number of global IPv4**
- **Has impact on new applications which need global**
  - **applications (particularly with security) or p-p**

13

19 June 2011   IWIN Workshop

Activities in Regions

- **Mature regions have much larger historic allocations**
  - **Hence have less urgency to move to IPv6**
- **Asia-Pacific and Africa have much worse problem**
  - **Many have major interest in large-scale growth**
  - **Particularly China, Japan, Korea had to move fast**
- **All three have large deployments**
  - **CERNET-2 all IPv6 with 10 Gbps**
  - **China says they will go all IPv6 by 2016**
  - **Japan has had WIDE development project since 1999**

14

19 June 2011   IWIN Workshop

Contents of Talk

- **The Internet and IPv4 and address depletion**
- ***What is IPv6, and its advantages***
- **How near is transition to IPv6?**
- **New Protocols and IPv6 drivers**
- **Some important applications that would benefit**
- **Training and 6DEPLOY**

15

19 June 2011   IWIN Workshop

What is IPv6

- **Successor to current IPv4 Internet Protocol**
  - **Under development since about 1993**
  - **Ratified as Standard by IETF around 2001**
- **Principal characteristics**
  - **Much more address space – 128 bits instead of 32**
  - **Mobile IP support mandatory (better than in IPv4)**
  - **IPSEC mandatory (could be done in IPv4)**
  - **Better auto-configuration**
  - **Better multicast**
  - **More space for flow-control options**
  - **More efficient processing of header options**

16

19 June 2011   IWIN Workshop

Some Newer Protocols

- **Since 2000, most new protocol work in IETF has concentrated on IPv6 versions**
  - Could have done IPv4, but often did not bother
- **Is reflected in many recent protocols**
  - Better mobility support (MIP6) – no triangular routing
  - Multicast now more deployable
  - Robust Header Compression (ROHC)
  - Low-power support for wireless (6LoWPAN
  - More support for Quality of Service (in Header, but not always implemented in routers yet)
- **Nevertheless, most concentrate on Address Space**

17

Why was it not adopted years ago?

- **Needed complete new suite of programs in each component of the infrastructure and terminal**
  - Virtually all the components are now in place
  - Mostly in dual-stack mode so that either version usable
- **Needed clear concept of how to do transition**
  - This will clear be done via dual-stack
  - Mechanisms for operational transition now defined
- **Needed technical and/or economic reason to move**
  - Killer applications only slowly emerging
  - Address space depletion put off by technical measures and less serious in North America and Europe
  - Considerable concerns of cost/benefit of transition – training, equipment, disruption

18

Contents of Talk

- **The Internet and IPv4 and address depletion**
- **What is IPv6, and its advantages**
- ***How near is transition to IPv6?***
- **New Protocols and IPv6 drivers**
- **Some important applications that would benefit**
- **Training and 6DEPLOY**

Current Status

- **Impact of address depletion imminent**
- **Major studies done on cost of transition**
  - **E.g. GSA, DoD in US**
- **Research activities 2000-2006 showed ease of putting dual-stack in the network core and terminals**
  - **DoD pilots and testbeds 2005-2007**
- **Most Research networks now dual-stack**
- **Terminal equipment often has IPv6 1st choice**
  - **Microsoft since Vista, IPv6 preferred, goes to IPv4 if needed**
  - **Mobile have IPv6 since v 4.1 of WIN-CE6, Symbian OS7.0 Linux and BSD have long had IPv6 standard**
  - **Most big providers move to dual stack**
    - **Though not all applications as complete (e.g. Cisco VoIP/CUCM)**
- **June 8 IPv6 day**

IPv6 Day

- **Big providers showed IPv6 readiness**
  - **Google, Yahoo, Akimai, Cisco**
- **Large User Organisations showed they were ready**
  - **E.g. NRENs, SPAWAR, some enterprise sites**
- **Traffic proportion IPv6 still low – typically 0.1%**
  - **But went up 60% on day dropping to 30% thereafter#**
- **However there were few real problems**
  - **Most solved on the day**
- **Expect steady migration of major services permanent**
  - **Also expect IPv6 week in February 12**

21

19 June 2011 IWIN Workshop

The Pioneers Advance Slowest

- **Europe and North America started Internet earliest**
  - **Have by far largest allocations, hence less need to move fast**
  - **Have large deployment, so reluctant to move fast**
  - **Have most advanced components, so loathe to move fast**
- **China, Japan, Korea have greatest need and incentive**
  - **Do not have same reserves of IPv4 address space**
  - **Have more capital to invest currently**
  - **Would like to leap-frog Europe and North America**
- **US main computer components well developed for IPv6**
- **Europe and N. America slow to move processes and applications to IPv6**

22

19 June 2011 IWIN Workshop

Stages of Adoption of IPv6

- **There have been many studies of the stages needed to transition to IPv6**
  - – **A good one will be published shortly by the ECC committee of CEPT**
- **The report outline the stages for public IPv6 transition:**
  - – **1 The core backbones must go dual-stack**
  - – **2 The ISP must embrace dual-stack working**
  - – **3 Content servers must become IPv6 accessible**
  - – **4 User equipment may become dual stack**
- **The report also analyses the progress of the different countries along this path**
  - – **In general, it is a long haul to get organisations transitioned**

23

19 June 2011   IWIN Workshop

US DoD Transition Good Case Study

- **2001 Electronics Board tasked to produce strategy**
- **2003 Came up with broad policy**
  - – **All new systems from 03 be IPv6 capable, IPv4 Interoperable**
  - – **Support testbed (NAVIPv6) in university**
  - – **Identify a at least 3 major projects that could be IPv6 Pilots**
  - – **Transition 2005 – 2007**
  - – **DISA manage and control all IPv6 address space for DoD**
- **Set up labs and testbeds**
  - – **With ever increasing functionality**
- **Set major standards for DoD**
- **Built database of accredited suppliers and applications**
  - – **Working closely with industry**

24

19 June 2011   IWIN Workshop

Japan More General Strategy

- **WIDE Project worked on IPv6 from 2000**
  - Strong involvement from industry
  - Director, Murai, moved to Prime Minister's Office
  - Built IPv6 infrastructure around 2000
  - KAME to provide IPv6 OS around 2000
  - Worked on mobile applications (and cars)
  - Equipped major building in Keio U for energy monitoring and conservation
- **Sony early research activity including 6NET**
  - 2004 stated all relevant future projects would be IPv6
  - Withdrew from effort on in games in 6NET to continue it in Japan
  - Games are p – p and need the IPv6 addresses

19 June 2011   IWIN Workshop

25

European Framework Research

- **Significant pilot network projects 2000 – 2005**
  - 2000-2003 6INIT (infrastructure), 6WINIT (mobile apps)
  - 2003-2005 Serious Pilots 6NET (network plus apps), EuroIX (Internet exchanges), Security
- **Training and Applications 2006 – 2009**
  - 2006-2010 6LINK, 6DISS, 6DEPLOY, 6CHOICE
  - 2007-2009 Civil Protection (U2010), 6Power, 6SAT
- **From 2010 no particular IPv6 Projects**
  - But assume that most projects will use IPv6 in their execution
- **Research Infrastructure GEANT dual stack**
  - Most European NRENs also dual stack
  - Very few universities have much IPv6

19 June 2011   IWIN Workshop

26

Many Actively Promoting IPv6

- **IPv6 Forum frequent Awareness Meetings**
  - **Many national IPv6 Task Forces**
  - **IPv6 Readiness Logos**
- **2008: European Commission IPv6 Action plan**
  - **Propose 25% users be able to connect with IPv6 by 2010**
  - **Proposes EC and EU e-Gov sites be enabled**
- **2009: 1st EU Agency provides IPv6 web access**
  - **European Network &Information Security Agency (ENISA)**
- **IPv6 EU Deployment Monitoring Survey**
  - **By TNO, GNKS Consult and RIPE**
  - **610 respondents, including government bodies, ISPs, other technology houses, and education**

19 June 2011   IWIN Workshop   27

Survey results: European IPv6 use

- **79% have or in process of getting IPv6 addresses**
  - **97% of educational institutes have IPv6 addresses**
- **17% using IPv6**
  - **8% of ISPs are using IPv6**
- **30% concerned about IPv4 depletion**
  - **Compared with 48% concerned outside the EU**
- **Why not deployed yet?**
  - **70% No business case**
  - **57% lack of user demand**

19 June 2011   IWIN Workshop   28

Contents of Talk

- **The Internet and IPv4 and address depletion**
- **What is IPv6, and its advantages**
- **How near is transition to IPv6?**
- ***New Protocols and IPv6 drivers***
- **Some important applications that would benefit**
- **Training and 6DEPLOY**

29

New Protocols

- **Survey indicates lack of interest or understanding of urgency**
- **Neither organisations nor user understand the impact of protocol progress over last eight years**
- **IETF has concentrated on IPv6 with new protocols**
  - **Many could be developed for IPv4, but have not been**
  - **Examples are improved 6LowPAN (low power protocols), ROHC (Robust Header Compression), MIP6 (mobile users), NEMO (mobile networks), MANEMO (Mobile ad hoc)**
- **Thus many of the future applications do not really have good IPv4 protocol support**

30

Future Driving Needs for IPv6

- **Know predicting future is a mug's game**
- **Mobile Important driver**
  - **IMS needs global access, agreed that it be IPv6**
  - **As VoIP goes mobile, needs many addresses, not IPv4**
- **Smart grids being developed globally**
  - **Needs many addresses**
- **All peer-peer traffic**
  - **Games, VoIP, Conferencing, Supplier push advertising**
- **Major interactive automobile services**
  - **Again problems of data push if private addresses**

19 June 2011   IWIN Workshop

31

Contents of Talk

- **The Internet and IPv4 and address depletion**
- **What is IPv6, and its advantages**
- **How near is transition to IPv6?**
- **New Protocols and IPv6 drivers**
- **_Some important applications that would benefit_**
- **Training and 6DEPLOY**

19 June 2011   IWIN Workshop

32

The Internet of Things

- **There is a general agreement that the _Internet of Things (IOT)_ will be the biggest future Internet User**
  - **Though nobody really knows how and what it is**
- **The EC has a major IOT programme**
  - **Which has world-wide international links**
- **At least the EC programme believes that while many end-devices will not be IP, wide-area communication must be Internet**
  - **They even believe it will have to be IPv6**
- **Several projects are investigating the role of IPv6**
  - **Including IOT6 which includes UCL**

33

19 June 2011   IWIN Workshop

IOT6

- **IOT6 is investigating how IPv6 addressing can be used to reach non-IP entities**
  - **Here the large address space and the local use of multicast may allow robust use of gateways**
- **Wide Area it will capitalise on the strengths of IPv6**
- **Its gateways will be mulit-level, and may be distributed – using IPv6 features**
- **We will investigate which information should be carried at the packet level**
  - **And which must be at the application level**
- **There will be a major practical demonstration in automating new buildings**

34

19 June 2011   IWIN Workshop

Smart [power] Grid

- **Smart Grids are being developed globally**
  - **Make grid more efficient – potential large cost savings**
    - US estimated $56- $112 Billion saving in 20 years
    - 2005: Italy - Telegestore project €2.1B – annual savings €500M per year!
- **2009: US Smart Grid Initiative - $8.1 Billion**
  - **40 Million smart meters**
  - **http://www.nist.gov/smartgrid**
  - **Smartgrid BoF at IETF76 in Japan, Nov 2009**
  - **Happening fast – standards to be ready by end 2010**
- **Large number of addresses => Need for IPv6**
  - **Could be done with IPv4 and private address spaces but would be much harder and constrain customers**

19 June 2011   IWIN Workshop

35

Europe moving to Smart Grid BUT

- **Electricity and Gas embracing remote meter reading**
  - **But motivation is narrow and industry interest dominated**
  - **Example is British Gas use with GPRS and without Internet**
- **If they go Internet at all, use heavily NATs with IPv4**
- **Customers may want their own LAN-based system that includes all utilities, local generation and local control**
  - **This will need IPv6 for addressing and protocol reasons**
  - **If working with NATed IPv4 utilities, will require complex G/w**
- **It is not clear how this will evolve in Europe**
  - **For instance British Gas is NOW introducing "smart meters" which will use GPRS to utility, and special ZIGBEE to customer**
- **We pray that new start-ups will obsolete the majors**

19 June 2011   IWIN Workshop

36

Emergency Communications

- **U-2010 showed applicability of IPv6**
  - Significant Luxembourg demo with fire, police & ambulance
- **Some of the conclusions of the EC IP**
  - Gateway to TETRA, but much better performance
  - Large-scale addressing of sensor networks
  - Capability of dealing with adhoc network
  - Ability to deal with security of sensor nets and media
  - Addressing size allows federation of different agencies on specific VPNs
  - Autoconfiguration allows easier set up of networks when infrastructur has been destroyed
  - Support for mobility in IPv6 particularly important

19 June 2011 IWIN Workshop

37

Emergency Communications Future

- **Some Europeans very interested in U-2010 results**
- **Most concerned at the size of present deployments**
  - Even IPv4 new to most
  - Strong pressure to deploy TETRA
  - Mobile operators interested in IPv6 mainly for IMS
- **This is in spite of clear advantages of IPv6 for this application**
- **Requires relevant authorities to look at transition questions in the light of current TETRA deployments**

19 June 2011 IWIN Workshop

38

Personal Communications

- **Few VoIP and Conferencing systems fully IPv6**
  - *Though with scale envisaged, IPv6 would be needed*
- **Some Open Source products already enabled**
  - *ISABEL, VIC/RAT, Linphone, SIP-Communicator*
    - **Though not all completely IPv6-tested yet**
  - *OPENSER and ASTERISK have open-source IPv6 versions*
- **Less Commercial products fully available, but e.g.**
  - *Cisco has product (with limited protocol support)*
  - *Tandberg is IPv6 ready (and is now Cisco)*
  - *There Is still very limited inter-vendor testing*

19 June 2011   IWIN Workshop

39

What should we do for Services

- **Ensure backbone networks are dual stack**
- **Ensure your main servers can run dual-stack**
  - *Web, file, message*
- **Ensure your local infrastructure has dual-stack capability**
  - *Running via tunnels to other islands if necessary*
- **Evaluate major software systems you use are IPv6-ready**
  - *Ensuring new procurements have dual stack upgrade clauses*
- **Ensure that terminal equipment is IPv6-ready**
- **Start running dual-stack in your organisation**
- **Start running some IPv6 services – like conferencing or web**
  - *Using tunnels if other infrastructure not ready*

19 June 2011   IWIN Workshop

40

What should we do for new Apps

- **Once you are running some sort of IPv6 infrastructure, it is worth exploring where advanced IPv6 features would help, e.g.**
  - Mobile applications or ones with ad hoc nodes, where MIPv6, NEMO or MANEMO will be useful
  - Peer-peer applications, where you will run out of address space
  - Large-scale monitoring applications, where both large address space and 6LoWPAN will help
  - Emergency situations where the address space helps automated VPN construction, auto-configuration helps and the built-in IPv6 are particularly helpful

41

19 June 2011   IWIN Workshop

Contents of Talk

- **The Internet and IPv4 and address depletion**
- **What is IPv6, and its advantages**
- **How near is transition to IPv6?**
- **New Protocols and IPv6 drivers**
- **Some important applications that would benefit**
- ***Training and 6DEPLOY***

42

19 June 2011   IWIN Workshop

Training

- **Clearly training is a major need**
  - **There are already many initiaiives**
- **Cisco Academy recently reviewed all its module**
  - **Now many consider IPv6**
- **ISOC, Regional Registeries and some commercial courses**
- **The 6DEPLOY series of IPv6 training projects**
  - **Provides one to three day courses each year**
    - **Mainly in emerging economies**
  - **Has strong practical component, with local and remote labs**
  - **provided by Cisco**

19 June 2011   IWIN Workshop

6DEPLOY Project

- **6DEPLOY is training Project**
  - **Has produced some 30 modules for training**
  - **Provides dozens 3-day training each year**
  - **Sometimes shorter**
  - **Has strong practical component using labs**
- **Cisco donated labs to project**
  - **By end of 2011 20 labs, 2 in this region**
- **All labs now have standard equipment**
- **Some act as Standard Labs**
  - **Common booking system, procedures, addressing**
- **Workshops can use labs locally, remotely, together**

19 June 2011   IWIN Workshop

Intentions of Project and Labs

- **The intention is to develop centres of expertise**
- **The labs are distributed regionally in order to encourage them to foster regional training**
  - **Are concentrating on emerging economies**
  - **Now mainly routers, with locally supplied PCs**
  - **soon also VoIP and Sensor nets**
- **Cisco provides hardware and software support**
- **6DEPLOY provides an initial training workshop**
- **The intention is that the institutions quickly become self-sufficient and do their own training**
- **It is hoped that the labs will all cooperate strongly**

19 June 2011   IWIN Workshop

45

New Features in Labs Planned

- **Server-based software routers**
  - **Based on Cloud Computing**
  - **Will co-operate with physical routers**
  - **Will be able to use physical or virtual routes locally or remotely**
  - **Some will be provided in remote labs, some will be provided as a cloud for Internet use**
- **Developing some new applications areas**
  - **VoIP, Conferencing, Sensor Neworks**
  - **Standard Labs may be equipped with these in the future**

19 June 2011   IWIN Workshop

46

Location of 6DEPLOY Laboratories

- **Cisco has established laboratories in four continents**
- **Europe**
  - Bulgaria (Sofia), Czech Republic (Prague), France (Paris), Hungary (Budapest), Slovenia (Ljubljana), Spain (Madrid),
- **Asia**
  - Georgia (Tbilisi), India (Bangalore), Kyrgyzstan (Bishkek),
- **Africa**
  - Cameroon (Yaounde), Egypt (Cairo), Ghana (Accra) Kenya (Nairobi), Mauritius (Ebene), Rwanda (Kigali)
- **America**
  - Columbia (Bogota),

47

19 June 2011 · IWIN Workshop

Some Closing Thoughts

- **The Internet is now all pervasive**
- **While it can no longer change in a revolutionary way, it will change hugely in an evolutionary one**
- **The next step must be IPv6**
  - Transition has been slow, but must now speed up
- **USA has led the provision of IPv6-based components**
- **However South and East Asia are leading the transition**
- **Some huge new applications *should* be based on IPv6**
  - A combination of blinkered and short-term thinking may prevent this happening in some regions
- **Training is vital everywhere for the transition**
  - special measures are needed in emerging countries

48

19 June 2011 · IWIN Workshop

Session 2:
Business models and Security
(Chair: Takuya Yoshihiro)

Application Push & Play – Proposal on Dynamic Execution Environment Combined with Personal Devices and Cloud Computing. -

Hidenobu Ito[*], Kazuaki Nimura[*] , Yousuke Nakamura[*], Akira Shiba[*] , and Nobutsugu Fujino[*]

[*]Fujitsu Laboratories Ltd.
4-1-1, Kamikodanaka, Nakahara-ku, Kanagawa, Japan
{itou.hidenobu, kazuaki.nimura, nkmr, shiba.akira, fujino}@ jp.fujitsu.com

*Abstract* – Currently mobile devices become essential tools in our convenient life. Especially in order to make effective use of smartphones, it is common that a user installs a lot of tiny applications in advance and flexibly uses them based on his/her circumstance. In such case, it is getting too larger to ignore how much time and cost a smartphone user spends to discover and install appropriate applications. In this paper, we propose a concept wherein desirable applications and data automatically descend from the Cloud and are executed on a device when a user is authorized to receive a service. Such applications are available only at the right moment and then go away when a user no longer requires them. As a proof of concept, we implemented the system on an Android smartphone and evaluated and confirmed that it can reduce installation steps and can work at good enough response. The proposed system would be able to reasonably help user's daily activity and provides a new user experience different from a conventional one.

Keywords: Mobile device, Cloud computing, Push services, Android

1 INTRODUCTION

A variety of personal devices such as smartphones, tablets, and personal computers are now appearing on the market. In addition, there are so many applications on such devices.

In terms of the usability of personal devices, it is not necessarily easy to start using them. Especially, there is the issue of the initial setup being troublesome and sometimes annoying. For example, those who have purchased a smartphone will waste much time finding and installing applications from an app store, etc. which they are likely to use frequently. Without any preinstalled applications, they would not be able to use it conveniently. They hence have to find and install applications before using it. In addition, installing too many applications on local device also makes them inconvenient because it would be difficult to identify an application they want to use on small screen of smartphone. In short, we would not like to spend much time to set up a smart phone because it has a short life cycle. Therefore, it would be good to have a system where a user can receive services with minimal preparation.

There are several articles about usage of applications for smartphones. Smartphone users use applications as important tools to manage information, tasks, works, and social relationships in their busy lives. Time management is one of the most important factors for them.

According to the report of SBE Council[7], mobile application discovery is a big problem for smartphone users.

On the other hands, there are several research results that show that the retention rate of mobile applications is not high. Ringcentral Survey[6] found that the retention rate in six months is 36%. Scott Kveton, CEO of Urban Airship, a mobile notifications provider, stated that there is only a 5% retention rate on free applications after 30 days.

Therefore, it is not always true that smartphone users can save time by utilizing mobile applications because they waste too much time for preparing seldom-used applications. In other words, reducing the time for discovery and installation of applications can give them a better mobile experience.

In this paper, we propose a concept of dynamic installation and execution which is named Application Push & Play (APnP) in order to resolve the above mentioned problem. In addition, we propose architecture for a smartphone to realize the concept. Then, we prototyped the system by utilizing web applications executable on local device and an Android smartphone. Finally, we evaluate the feasibility and efficacy with the results of experiment.

2 RELATED WORK

Apple Inc. provides the App Store on iTunes[4] to distribute applications for iPhones. An iPhone user can download applications via iTunes on a PC, and synchronize and install them on an iPhone through a USB connection. A user can also download and install applications directly on a iPhone. Google Inc. provides the Android Market[3] for Android smartphones as well. The advantage of the Android Market is that you can download and install applications from a PC to any android smartphones you own, even if you have more than one device. On the other hands, Apple has recently started to provide iCloud to synchronize all applications between devices.

In terms of applications for mobile devices, Jason Grigsby[5] presented comparison among three types of applications which is Native, Web, and Hybrid. 'Hybrid applications' is defined as applications which are executable on local devices and are described by HTML, CSS and JavaScript. Though they look like web applications, they can run in a dedicated runtime environment on a local device. As of Web applications, HTML5[1] specifies a disk cache which can let web applications run on a local device even when the network is disconnected.

3 PROPOSED ARCHITECTURE

As described in the introduction, to reduce the time for discovery and installation of applications is a big issue. Additionally, it should be processed without a user's operations in order to save not only time but also labor. In this section, we define the requirements and propose the architecture to fulfill them.

3.1 Requirements

The requirements to solve the problem are as follows:

(R1) Applications exist on a personal device only when a user needs them.
(R2) A user needs not explicitly install/uninstall applications.
(R3) A set of suitable applications are automatically to be identified by a user's context.

The best way to resolve the issue of application discovery and installation is that a user does not care about application. All a user wants to do is to deal with contents. It does not matter for a user what kind of application can operate the contents. Thus, the requirement is that not a user but a system can identify applications which a user needs and can deliver them to a personal device just in time. (R3) is equivalent to the requirement of identification. The requirement of delivery consists of adequate notification (R1) and automatic installation (R2).

3.2 Architecture

We have come up with a concept of Application Push & Play shown as **Figure 1** in order to fulfill the above requirements.
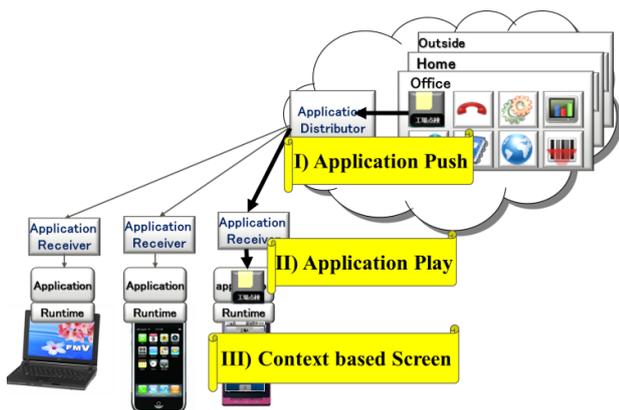

Figure 1: Concept of APnP

The architecture consists of the following features:

(A1) Application Push
(A2) Application Play
(A3) Context-based Screen

Our approach assumes that all applications are in the Cloud environment and are downloadable and executable on local devices whether network connections are available or not.

'Application Push' (A1) handles distribution of applications and data based on a user's context which might be derived from sensor information and user behaviors, etc. Once the system has detected a user's state as likely to use a service, a particular application in the Cloud would be injected into the local storage of the user's device.

'Application Play' (A2) deals with automatic installation and execution of applications injected by 'Application Push'.

'Context-based Screen' (A3) provides a user with a simple and comprehensible GUI. It is a dynamically morphing home screen in accordance with the place and time of user.

To meet the requirement (R1), the system should have a method of real-time notification to deliver application timely. In general, there are two methods about notification: "polling" and "push". Polling method is that a client checks if there is an event to a server periodically. Typical example is an e-mail system common in a PC. Push method is that a server notifies to a client when an event occurs. Typical example is a SMS system common in a cell phone. In terms of power consumption, polling method is not favorable to a smartphone. That is why push notification of application (A1) is an important feature. It includes push server and push handler as a client.

To meet the requirement (R2), the system should have a method of automatic installation (A2). It includes management functionality of downloading, storing data, and launching applications based on commands in a pushed message.

To meet the requirement (R3), the system should have a method for a user to reach an application identified at a glance under particular context. The definition of user's context is also important. In this architecture, however, how to define and make a decision about a user's context is out of scope. It can be placed in the Cloud and collect any information such as sensor data about a user. A simple example of context engine is location-based or schedule-based system. Typically, a task server includes a context engine and invokes notification via a push server. Our focus is to provide a user-friendly interface as a home screen (A3) by utilizing an existing context engine.

Figure 2 shows an APnP Framework described as the abobe.
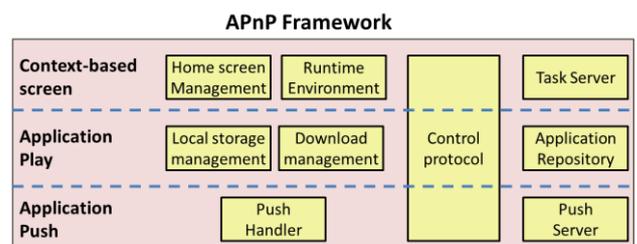
Figure 2: APnP Framework

4 PROTOTYPE IMPLEMENTATION

Figure 3 shows a prototype system to prove the feasibility of our concept.
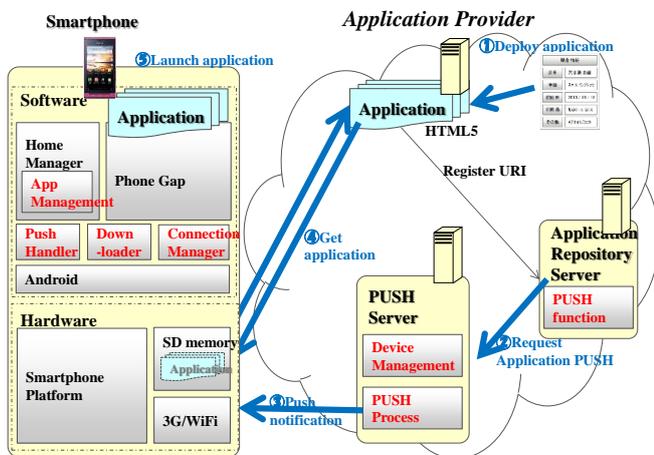

Figure 3: Implementation of the proposed system

4.1 Components

Proposed system consists of the following components:

'Application Repository Server': This stores universal resource indicators (URIs) of applications. It picks an appropriate URI up on order from 'Task Management,' and then commands 'Push Server' to send a PUSH message with the URI.

'Application Provider': This stores an actual application and provides an interface to send it by request.

'Task Management': This manages the user's context on a basis of calculating data from sensor information. Once it evaluates the timing to send a particular application to a particular user, it issues a command to 'Application Repository Server' to send the application.

'Push Server': This constructs PUSH messages and sends them to client devices. On receiving a command from 'Application Repository Server' it resolves an appropriate client device to send a PUSH message to with an application URI. Google already provided the PUSH mechanism called Android Cloud to Device Messaging Framework[2]. However, it is not suitable for business users because it requires Google Account. Therefore, we implemented an dedicated PUSH Framework.

'Push Handler': This is a background process working on a client device to receive a PUSH message from 'Push Server'. It keeps a TCP connection to 'Push Server' and can communicate with 'Push Server' as long as physical network connectivity by 3G or WiFi is maintained. Once it receives a PUSH message and it finds an APnP command in the message after analyzing, it hands the command to 'Home Manager'. In this implementation, 'Push Handler' is realized as an android application.

'Downloader': This is a client module to download applications from 'Application Provider' to SD memory in the smartphone by request from 'Home Manager'. In this implementation, 'Downloader' is realized as a library module which is included in the 'Home Manager'.

'Home Manager': This displays a home screen. It has three functionalities. First, it is an application launcher to execute applications. Second, it performs management of icons to be shown on the home screen. The home screen is dynamically changeable on the basis of APnP commands from 'Push Handler'. It displays only icons necessary in context to provide a simple user interface. Finally, it orders 'Downloader' to download applications. In this implementation, 'Home Manager' is realized as a kind of Android application.

'PhoneGap'[8]: This is a runtime environment to execute applications stored on SD memory in the smartphone.

4.2 Work flow

A typical work flow is as follows:
1. Deploy applications
   Developers distribute their packaged applications via an application server provided by an application provider. After that, they register the URI of the application to the application repository server. IT administrators might be able to do for enterprise usage.
2. Request Application Push
   Once context engine which knows user's behavior detects an event where a user wants to use the application, it creates a command message, and then requests a notification to the target device via the push server.
3. Push notification
   After receiving the request from application repository server, a push server sends a push message which includes APnP command to download and invoke the application.
4. Get application
   When the push handler inside the smartphone receives the pushed message, it analyzes the APnP command and then download zipped resources of the application by utilizing HTTP protocol. After complete to download, it stores them to the SD card which is accessible for the home manager and PhoneGap.
5. Launch application
   Finally, the downloader sends the Intent to the home manager to notify that download is completed. The home manager changes the screen and launch the application by invoking PhoneGap.

4.3 APnP Commands

APnP commands in a push message consist of the following elements:

Table 1: A list of APnP Commands

Command Name	Description
APCTL	Operations by Downloder, such as download an application, update a part of application, delete an application, etc.
HMCTL	Operations by Home Manager, which is display options how it shows notifications to a user.
HMOP	Option values which Home Manager use, such as context information, delayed time to execute application, etc.
URI	The URI of applications.
RESOUCES	The resource name of application to be downloaded.

4.4 Packaged application

An application is packaged as a set of resource files, a list of resources and a manifest file.

Figure 4 is an example of a list of resource files. All files are described as a plain text.

```
Index.html
pageA.html
pageB.html
webapp-manifest.json
walkingApp.manifest
images/walking.png
images/title.jpg
```

Figure 4: Example of a list of resource files.

In this example, HTML files ('*.html') and image files ('walking.png', 'title.jpg') are actual contents of the web application. These set of files are minimum resources in a common web application. Additional resources specific in this system are 'walkingApp.manifest' and 'webapp-manifest.json'. A 'walkingApp.manifest' is a name of this file. it is specified in HTML5[1] to realize local execution. The resources described in this file are cached in local storage, and can be executed without communication to a web server. A 'webapp-manifest.json' is a manifest file to describe properties of the application. Figure 5 is an example of 'webapp-manifest.json'.

```
{
  "appURI":
"http://www.example.com/walkingApp/",
  "appName": "walking",
  "description": "This application recommends a
good walking place.",
  "creator": "fujitsu_healthcare@example.com",
  "version": "1.0 ",
  "manifest": "walkingApp.manifest",
  "icon": "images/walking.png ",
  "toppage": "index.html"
}
```

Figure 5: Example of a manifest file.

An application package is uniquely identified by an 'appURI'. A Down loader refers to the appURL to download the application and deploys downloaded files on the local storage as the same folder tree. For example, files downloaded from http://www.example.com/walkingApp/ are placed on /sdcard/apps/www.example.com/walkingApp/. An 'appName' is a name of application. A 'description' is a detail explanation of application. A 'creator' is contact address of developer. 'version' is a version number of application. A 'manifest' is a name of a HTML5 manifest file. A downloader can get where each resource is located. An 'icon' is an image file to show icon on a home screen. A 'toppage' is an entry page to be displayed first after launching application.

5 EVALUATION

5.1 System evaluation

In order to evaluate the feasibility of the proposed system, each component shown in Section 4 is operated on the following hardware:

Fujitsu PRIMERGY TX100: This is an IA server to run the Application Repository Server, PUSH Server and Applications. The specifications are Intel Xeon Processor E3120 3.16GHz for CPU, 8GB DDR2 memory, 1TB SATA HDD and 1000Base-T for NIC. The operating system is Cent OS 5.5

Fujitsu T-01C: This is a smartphone with Android 2.1 to run client-side modules, such as Push Handler, Downloader, Home Manager, and PhoneGap as a runtime environment.

Google Nexus S: This is a smartphone with Android 2.3 as well as T-01C.

Figure 6 shows practical example of screen transition.

External Memory	2GB Micro SD	Use a partition of internal memory
OS	Android 2.1	Android 2.3

Four kinds of implementation are determined by combination of how to download and store an application and whether an application is compressed or not. The differences of each implementation are as follows:

I) Download a zip-compressed file and store it in SD memory. After that, unzip it and store the unzipped files.
II) Download a zip-compressed file and store the unzipped files in SD memory by utilizing ZipInputStream class without storing the zipped file.
III) Download an uncompressed file packed by tar tool and store it in SD memory. After that, unpack it and store the unpacked files.
IV) Download the uncompressed file packed by tar too and store the unpacked files in SD memory by utilizing TarInputStream class without storing the packed file.

At this point, we can make the following hypothesis:

Hypothesis:
The best methods is that an application package is compressed and is stored it directly as unzipped files to a local storage.

To prove the hypothesis, we should confirm the followings:

A) Total throughput in Case-II and Case-IV are better than in Case-I and Case-III respectively.
B) Total throughput in Case-I and Case-II are better than in Case-III and Case-IV respectively.

It is expected that (A) and (B) is true if CPU is sufficiently fast. The reason why (A) should be true is because two times of file I/O operations occurs in Case-I and Case-III, whereas only one time of file I/O occurs in Case-II and Case-IV. The reason why (B) should be true is because the increase in data traffic has a large impact on the total performance compared reduction of the load on uncompressing process.

We measured an actual time spent from the beginning of downloading till the end of storing an application package for each case. In Case-I and Case-III, we also measured a download time and an unpacked (or uncompressed) time. The sum of both times should be a total time.

The application package to be downloaded is a 167KB file zip-compressed by 25 files and 5 folders. The original total size before compressed is 655KB. The Compression rate is 27%. The application is downloaded through the WiFi connection.

The picture in the left is a snapshot of home screen. In this example, it displays only icons suitable to outside when a user stays out. Example of context is outside, home, work, etc.

The picture in the middle is a snapshot of notification that a business application comes in the meeting room. In this case, the home screen has been changed from "Outside" to "Meeting room" as soon as a user came into the company's meeting room. After that, a new icon of the application appears in the center of screen, and then the small icon is deployed on the home screen. It is optional just to display icons on the home screen without invocations of an application.

The picture in the right is a snapshot of running the notified application. After notification, the system can automatically execute the application without a user's operation.

After a user finish to use the application, it might disappear from the device if a user no longer needs it.

It provides a simple user interface to a user, and helps to discover application icons in the particular context.

5.2 Performance evaluation

5.2.1. Test Environment

In order to improve user experience, a time from when a user notices a notification till when he/she starts to operate is very important. As of the proposed system, a time from receiving PUSH message till appearing an icon or till invoking application would be bottleneck to provide a user sufficient response rate. Therefore, we evaluated the difference of performance for four kinds of implementations on both T-01C and Nexus S. The detail specifications of each smartphone are shown as Table 2.

Table 2: Specifications of T-01C and Nexus S

Model	T-01C	Nexus S
Manufacturer	Toshiba	Google/HTC
CPU	QSD8250 1GHz	S5PC110 1GHz
Internal Memory	512MB RAM	16GB iNAND

Figure 6: Example of Context-based Home Screen

5.2.2.　Test results

Table 3 shows the results of the measurement. Each time in Table 3 is average of 10 times measured values.

Table 3: Measured results

Item #	Device	Case #	Total time (msec)	Download time (msec)	Unpack time (msec)
N-I	Nexus S	I	203.2	110.2	93
N-II	Nexus S	II	190.8	-	-
N-III	Nexus S	III	434.4	264.6	169.8
N-IV	Nexus S	IV	402	-	-
T-I	T-01C	I	817.6	197.9	619.7
T-II	T-01C	II	653	-	-
T-III	T-01C	III	1318.6	524.6	794
T-IV	T-01C	IV	861.4	-	-

Figure 7 shows the comparison of the elapsed times for each test case by histogram.
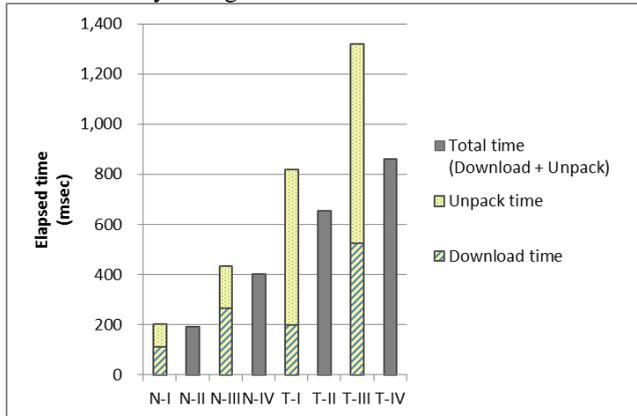
Figure 7: Comparison of the elapsed times.

Figure 8 shows the comparison of the throughput calculated from the result in **Table 3**.
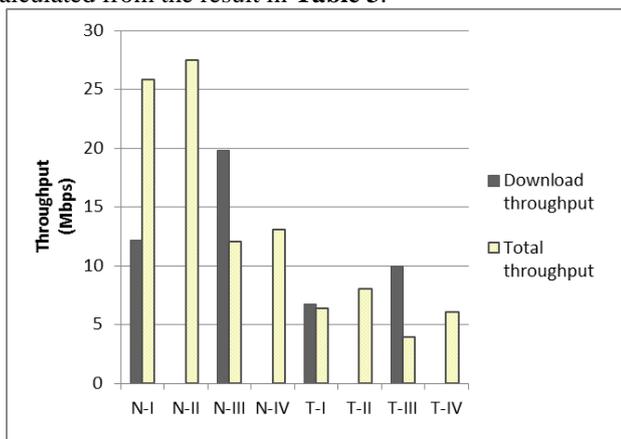
Figure 8: Comparison of the throughputs.

5.2.3.　Consideration about experimental results

Both histograms tell the following fasts:

i). Nexus S is approximately twice faster than T-01C in all cases.
ii). The fastest implementation is Case-II on both Nexus S and T-01C.
iii). There is little difference of the time between in Case-I and in Case-II or in Case-III and in Case-IV on Nexus S, whereas there is a significant difference on T-01C.
iv). On Nexus S, the cases of using a compressed application package such as Case-I and Case-II are approximately twice faster than the cases without compression of application package such as Case-III and Case-IV, whereas on T-01C, there is little difference of the total throughput between in Case-I and in Case-IV.

The reason why Nexus S is faster than T-01C is due to difference of the storage device between them. The application is stored in an external memory. Micro SD memory used on T-01C is significantly slower than NAND flash memory used on Nexus S. Therefore the difference of I/O performance has large impact for total throughput on the proposed system. The above Fact-iv is due to the same reason.

Even throughput of downloading on Nexus S is faster than on T-10C as well. It is presumed by difference of OS version. Android 2.3 is faster than Android 2.1 in general because JIT (Just-In-Compiler) is available on Android 2.3. However, T-01C can also have enough throughputs because even the worst time in total is less than a second if the application is zip-compressed.

In addition, the results of the case without compression of application package were worse than expected. It is natural that data traffic is getting more, download time is increasing. However, there was difference for unpack operations though there was supposed to be little difference of unpack operations. It is because performance of TarInputStream class on Android was too bad. At this time, an application package should be compressed.

To summarize, the experimental results satisfied both of the condition (A) and (B) described in 5.2.1. Therefore, we were able to confirm that the hypothesis in 5.2.1 is true.

As a result of experiment by utilizing an actual business web application, it was confirmed that application delivery was notified to the smartphone within a few seconds after an event had been issued, and then an application icon was dynamically added to the home screen, and the application was invoked shortly after.

6   USE CASE

In this section, we present an effective use case by utilizing an APnP system.

To simplify the user interface further, even the home screen would be removed. That is to say, an application is directly executed without displaying a home screen once it is distributed to our devices.
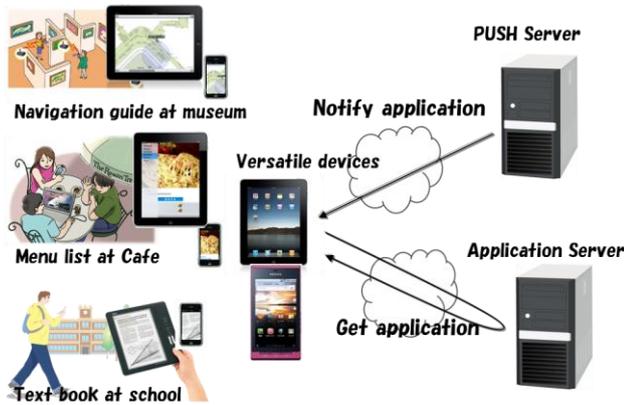
Figure 9: Example of service.

Figure 9 shows an example service. A user's device can be transformed into textbook while he/she is at school. At other times, it can be transformed into a navigational guide while he/she is in a museum. In such cases, a personal device with our system can behave as if it were a dedicated device, and can provide him/her with a rather simple interface. It might be more useful for a tablet device than for a smartphone.

7   CONCLUSION

We proposed a concept of APnP and prototyped a system where appropriate applications are automatically distributed and executed. Such applications are made available only at the right moment and then disappear when a user no longer requires them. We evaluated and confirmed that it can reduce installation steps and work with sufficient response. In conclusion, the proposed system can reasonably help in a user's daily activity and provides a new user experience different from the conventional one.

REFERENCES

[1] HTML5, http://dev.w3.org/html5/spec/Overview.html.
[2] Android Cloud to Device Messaging Framework, http://code.google.com/android/c2dm/index.html.
[3] Google Android Market, https://market.android.com/.
[4] Apple iTunes, http://www.apple.com/itunes/.
[5] Jason Grigsby, Native vs Web vs Hybrid Mobile Development Choices, http://assets.en.oreilly.com/1/event/34/Native%20Apps%20vs_%20Mobile%20Web%20vs_%20Hybrid%20Apps_%20Mobile%20Development%20Choices%20Presentation.pdf (2010)
[6] Ringcentral Survey, Small Business Professionals Admit to Smartphone Addiction, http://blog.ringcentral.com/tag/ringcentral-survey. (2011)
[7] SBE Council, Saving Time and Money with Mobile Apps – a small business "app"ortunity –, http://www.sbecouncil.org/uploads/Mobile%20APP%20Final%20Report%20SBE%20Council.pdf, (2011)
[8] PhoneGap, http://www.phonegap.com/.
[9] T-01C, http://www.nttdocomo.co.jp/english/product/foma/smart_phone/t01c/index.html
[10] Nexus S, http://www.google.com/nexus/

Experimental transformations between Business Process and SOA models

Akira Tanaka* and Osamu Takahashi**

*School of System Information Science, Future University Hakodate, Japan
**Department of Media Architecture, Future University Hakodate, Japan
{g3109002, osamu}@fun.ac.jp

Abstract - in designing enterprise IT systems, two major architectural styles exist today: process-oriented and service-oriented architectures. Either one of them can be used to define behavioral aspects of the business specifications. In reality, a process can make use of various services, and a business service can be implemented as a process. This duality applies to such technology as BPMN and SOA. In RM-ODP standard, however, both are part of a standard viewpoint language, and they complement with each other. In this paper, using a textual domain specific language and a tool supporting it to capture the essence of those modeling languages, we examine the relationship between process-based specifications and service-based specifications for a typical buy-sell-ship business process or collaboration. Architectural comparison is done by examining a model transformation of process to service, and service to process. The difference of the two types of model and the implication of the results are discussed.

Keywords: business process; service-oriented architecture; enterprise architecture; RM-ODP; model transformation.

1 INTRODUCTION

Enterprise architecture is widely used as a way to describe overall architecture of enterprise systems. There are many approaches to define enterprise architecture. For instance, to describe everything from hardware components to business strategies, a matrix of concerns and perspectives is used in Zachman Framework [1], or a set of customized architectures such as application architecture to capture different aspects is used in Federal Enterprise Architecture [2] initiatives. There are also requirements to harmonize businesses and ITs. In this context, the importance of business specifications, which describes "what to achieve" at business level, is being recognized.

There are a variety of approaches to define business specifications: business process oriented approach that uses BPMN [3], business rule oriented approach that uses SBVR [4], business events based approach such as EDA and CEP [5], business service oriented approach such as SOA [10], and many more. Among them, two major architecture styles exist, which are process-oriented and service-oriented architectures. They are mainly usaed to define behavioral aspects of the business specifications.

According to BPMN specification, "business process is a defined set of business activities that represent the steps required to achieve a business objective." This implies, when top down design approaches are taken, a business objective is set first, and a process or a set of processes is/are defined to achieve the objective. The IT system will be designed to provide necessary functionalities to the defined steps in the process. In this approach, interested behaviors or interactions exist between steps and among all the participants.

According to SOA Reference Model, "A service is a mechanism to enable access to one or more capabilities, where the access is provided using a prescribed interface and is exercised consistent with constraints and policies as specified by the service description." In this approach, most application elements will be implemented as services so that client software can find and consume necessary services to achieve its goal. The interactions in this model are between consumer and provider of the service. Orchestration of the services is not within the scope of this architecture.

Both styles are used as foundational architectures when developing enterprise IT systems. It should be noted, however, that a process can be decomposed into steps, each of which may consume services. And, a service can be implemented as a process. This duality applies to business systems designed based on current technologies such as UML [6], BPMN and SOA.

The issue we have is "are they really different, or are they different sides of the same coin?" In other words, are they only different in architecture styles and equivalent in capabilities? And if so, how can we measure the equivalence? In this paper, we will examine this issue using modeling technique including UML, Domain Specific Language (DSL [8, 9]), and model transformations. If they are essentially equivalent, there should be correspondence between them, that is, a service based model should be able to be transformed into a process based model, and a process based model should be able to be transformed into a service based model.

We will first look at business process oriented approach and examine how much it can be mapped to SOA approach. We will then look at service based approach and examine how much it can be mapped to process oriented models. In doing this, we will use DSL and model to text transformation tool.

2 BUSINESS PROCESSES

Business specifications are usually the most examined specification in Enterprise Architectures, since most business users, in addition to technology providers, would need to review it to see if it correctly captures the business requirements. When top down approach is taken, it usually starts with analysis of business environment, establish a new goal and strategies, design business processes to achieve the

goals. There are multiple choices in diagramming business processes. UML Activity Diagram and Business Process Modeling and Notations (BPMN) are the most used ones. In this paper, we use ODP ([7] see IV) Process diagrams that is a slight extension to UML Activity Diagram, since it has necessary characteristics to do the experiment.

The diagram on the right side of this page (Figure 1) shows a sample purchasing business process among three parties, buyer, seller, and shipper in ODP Process Diagram. Each lane represents role and behavior of the party, and each step is represented as an Action internal or external to the lane. In case of external Action, the control flow crosses the lane with or without artifact passing. Artifact, represented as ObjectNode, is there to capture necessary business information to be passed. There are split/merge used to control, i.e. to create and conclude, parallel activities. This is almost the basic Activity Diagram except for applied ODP stereotypes. The dotted lines are there to show logical grouping of the steps that have certain meaning in the application, e.g. placing an order. Note that these dotted lines are not part of standard notation, and they should be read as additional comment. Although it is possible to group steps using sub-process, it does not provide improvement in readability of the process, and therefore we did not take that approach.

In summary, this process diagram shows participants of the purchasing process and a collection of necessary steps in a prescribed manner leading to the objective. IT systems will be designed to support some portions of the steps. This style is effective when an IT system is to be built against pre-defined business processes (i.e. what needs to be done in what order).

3 SERVICE OR BUSINESS SERVICE

The definition of term "Service" in SOA is still under open discussion. However, there is one in OASIS's SOA Reference Model [10], which is "A service is a mechanism to enable access to one or more capabilities, where the access is provided using a prescribed interface and is exercised consistent with constraints and policies as specified by the service description." OMG's SoaML [11], which is a UML Profile for Service oriented architecture Modeling Language, is a standard to describe SOA based models. A slight modification of above definition was used there, which says "A service is value delivered to another through a well-defined interface and available to a community (which may be the general public). A service results in work provided to one by another." In this UML Profile, various SOA concepts are defined. For instance, a community is defined as a place for participants consume/provide services to each other. For each pair of participants they have service contract that govern the behavior aspects when consuming/providing services. As for diagramming, it mainly uses UML Collaboration Diagram, Class Diagram and Component Diagram.

If compared with the previous process model, participant can be considered as role. The diagram on the right side of this page (Figure 2) shows a sample high level Service

Architecture using SoaML to represent Buy-Sell-Ship collaboration.

Note that there are other kinds of diagrams associated with this high level description. For instance, Service Contract Diagram contains two roles, consumer and provider, with a sequence diagram specifying service message exchanges when provided services are consumed.
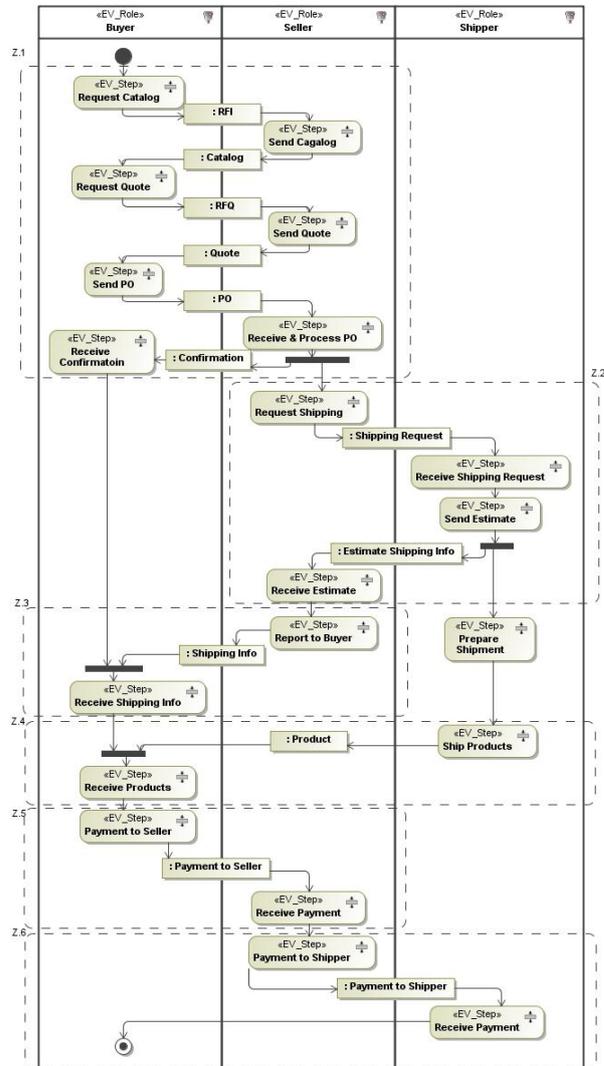

Figure 1 Sample ODP Purchasing Business Process
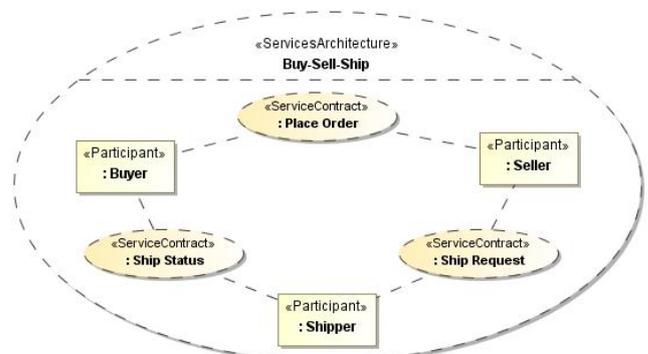
Figure 2 Sample Service Architecture

In summary, this service diagram shows participants of the collaboration and service contracts, which includes interface

and behavior definition, between service consumers and service providers. Note that there is no "steps taking place in a prescribed manner" defined. IT systems will be designed to support service providers and consumers. This style is effective when an IT system is to be built using existing or to be built services, such as newly developed internal services, wrapped legacy functions, or external services using web services, with a flexibility of clients' choice of the services.

4   RM-ODP

RM-ODP stands for Reference Model for Open Distributed Processing, which is a family of international standards for developing standards for open distributed processing systems. It is a set of reference models, and it also has UML Profile standard to represent its concepts using UML tools. This standard is used as an open Enterprise Architecture, and we use this standard with associated UML Profile (e.g. in Figure 1) to show something not biased to specific process modeling notations. It defines five standard viewpoints, but we will use or refer to only three of them in this paper: Enterprise, Information, and Computational viewpoints.

5   DOMAIN SPECIFIC LANGUAGE (DSL)

According to Domain-Specific Languages [8], domain – specific language is defined as "a computer programming language of limited expressiveness focused on a particular domain." DSL could be graphical or textual, could be internal (designed based on general purpose language) or external (having no specific host language). In this paper, textual and external DSLs for process oriented modeling (ProcessDSL) and service oriented modeling (ServiceDSL) are described and used.

6   PROCESS TO SERVICE MODEL TRANSFORMATION

Model transformation achieves one source model described in one specific language to be converted into target model described in other specific language, without violating rules for those languages. Typical example is to transform UML Class Diagram to Relational Table. Four layer meta architecture is usually used to explain the mechanism. As a standard, OMG's MOF/QVT [12] is the best known one. As open source projects, widely known examples are eclipse ATL [13, 14] and QVT.

The next table shows metamodel and main elements of Process models and Service models related to our sample process.

Table 1 Metamodel Elements

	Business process	Service
Metamodel	Process related part of RM-ODP Enterprise Viewpoint metamodel	SoaML metamodel
Main Elements	Object, Role, Process, Step, Action,	Participant, Service Contract,
	Activity, Artifact, Interaction	Service Architecture, Service Interface, Service Choreography, Message, UML Collaboration/Component

The following is a summary of how service metamodel element could be created with given process metamodel elements.

A. Participant/UML Component

A Participant (actually a Type) is equivalent to a Role in process. Since Participant deals with computation, a computational object with the same name should be introduced or assumed in the Process model side.

B. Service Contract

Service Contract can be considered as a concept representing interaction or behavior between two roles. ODP's Interaction is the closest concept, but it is not really a part of process modeling. Service Contract uses Sequence Diagram to represent the behavioral aspects.

C. Service Architecture/UML Collaboration

Service Architecture's closest concept is a collection of Interactions among all the involved roles.

D. Service Interface

The concepts like Interface, Operation, and Signature belong to Computational Object in ODP. If we could assume the existence of Computational Object with the same name as Participant, they are the corresponding elements.

E. Service Choreography

Service Choreography defines ordering of service messages between service consumer and provider. This can correspond to an ordered sequence of ObjectFlow involving the corresponding two roles in the process model.

F. Message

Information viewpoint of RM-ODP is the viewpoint where all the concerns on information within the system are defined. However, in Enterprise and Computational viewpoint, there is a need for information model and they are created based on the one defined in Information viewpoint. The Message data types are a collection of data types and structure of data elements visible in Enterprise viewpoint, and therefore those should be prepared as a part of process model.

From above, we can observe that when transforming a process to a service, the process needs to be decomposed into an ordered set of two party interactions, and interaction should be brought into a process model.

Figure 3 is a sample ODP Interaction Diagram showing Buyer as an initiator of the interaction, Seller as a responder of the interaction, and the references to various artifacts that are actually electronic document. The behavior of both sides is described using UML State Machine, which is different from the case with Service Contract.

When considering the model transformation, this Interaction Diagram could act as an intermediate from process to service, meaning that we can transform a process model to interaction model, and then transform it to service model.
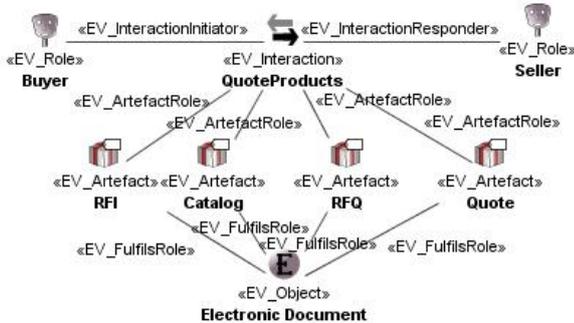

Figure 3 ODP Interaction Diagram

7 SERVICE TO PROCESS MODEL TRANSFORMATION

This section considers the reverse of the previous section, which is about transformation from service model to process model. Service model here is a model based on SoaML.

A. Object

SoaML's instance of Participant can be considered as Enterprise Object in a limited sense, since it also consumes and provides services like Computational Object. Message is a good source for defining Information Object.

B. Role

Participant is almost the equivalent to role. At the same time, Participant can be understood as Computational Object, which takes all the interface information from the Participant.

C. Process/Activity

Service Choreography specifies dynamic binary relationship of the behavior, and can be used to construct a portion of a Process. However, constructing whole process is not possible, since there is no orchestration information is available for construction of the whole process, including where to start the process.

D. Step/Action

Step/Action means action execution, which is defined in Service Contract and Service Choreography. However, it is not possible to generate Steps/Actions which do not correspond to interactions such as internal actions.

E. Artefact

Message is the only element to map to Artefact.

F. Interaction

Service Contract corresponds to Interaction where Consumer side corresponds to Initiator side and Provider side corresponds to Responder side.

8 PROPOSED MODEL TRANSFORMATION METHOD USING TEXUAL DSLS

When we refer to "model transformation," it usually means transforming models created with UML tools or some other specific tool such as BPMN tools (graphical tools). Those models are actually saved as text file, for instance as a form of XMI [15] or XML, and then transformation logics are applied to it. However, UML itself is a complex specification and that is reflected in XMI. A simpler way to experiment some modeling issues without involving too much complexity was needed, and that was the reason we took DSL.

We will now explain a method of model transformation using textual DSL. With textual DSL such as the one developed with eclipse/Xtext [16], a grammar is first defined, and the grammar based model editor is generated so that user can create his/her own model based on the DSL. Suppose you have two textual DSLs: ProcessDSL for process modeling and ServiceDSL for service modeling. If you define a grammar for ProcessDSL, you get the ProcessDSL editor. The same is true for ServiceDSL. Once a process model is defined, a template is applied to the model to generate output text (e.g. source code or XML file). Here, it is possible to design a template to generate text, which has a structure that ServiceDSL editor imposes. This is not always possible, since the source model may not contain necessary information to transform to. But, if it did, this model to text transformation works as a model to model transformation. This is the basic idea we used for Process to Service and Service to Process model transformations.

In order to achieve this, we have created above DSLs as simple textual DSLs to capture core concepts of ODP Process Diagram (or UML Activity Diagram) and SoaML Diagram respectively. They are simple, because not all concepts are used and some complex concepts were simplified to some extent in the grammar.

The tool used is eclipse/Xtext and its integrated model to text transformation engine Xpand/Xtend. But, with any other textual DSL tooling, such as Spoofax [17] or MPS [18], this can be done in the similar way.

Figure 4 shows a portion of the Process DSL grammar. Using generated DSL editor, a process model in this ProcessDSL is created (Figure 5), which is done by typing, not by applying model to text transformation from the process model (Figure 1). The last step is to define a template to generate text, which is explained in the next page. The Xtext grammar files and sample models in this paper are published on the following web site [19].

We also created InteractionDSL based on ODP concept of Interaction, and a sample model is shown in Figure 6. With this InteractionDSL, multiple artifacts are allowed to be included. However, in our transformation rule explained earlier, only one artifact per interaction is allowed.

```
Model:
    processes+=Process*;
Process:
    'process' name=ID '{'
        lanes+=Lane*
    '}' ;
Lane:
    'role' name=ID '{'
        processFlows+=ProcessFlow*
    '}' ;
ProcessFlow:
    InitialNode | EndNode | Step | Artifact | Para
InitialNode: 'start' name=FQN ;
EndNode: 'end' name=FQN;
Step: 'step' name=FQN ('{'
        ('description' description=STRING)?
        fromNode=FromNode
        toNode=ToNode
    '}')? ;
Artifact: 'artifact' name=FQN ('{'
        ('description' description=STRING)?
        fromNode=FromNode
        toNode=ToNode
    '}')? ;
ParallelSplit: 'parallel' 'split' name=FQN ('{'
        ('description' description=STRING)?
        fromNode=FromNode
        toNodes+=ToNode*
    '}')? ;
```

Figure 4 Partial ProceeDSL grammar

```
process Buy_Sell_Ship {
    role buyer {
        start StartOfPurchasingProcess
        step RequestForInformation {
            description "Gathering Product Information"
            from StartOfPurchasingProcess
            to seller.ReceiveRFI with Artifacts.RFI
        }
        step ReceiveCatalog {□
        conditional split PurchaseFromCatalog {□
        step SendPO {□
        step ReceivePOConfirmation {□
        parallel merge MergeWithInformation {□
        step ReceiveShipmentInformation {□
        parallel merge MergeWithProductsDelivery {□
        step ReceiveProducts {□
        step PaymentToSeller {□
        end EndOfPurchasingProcess
    }
    role seller {□
    role shipper {□
    role Artifacts {
        artifact RFI {□
        artifact Catalog {□
        artifact PO {□
        artifact POConfirmation {□
        artifact ShippingRequest {□
        artifact ShippingCostEstimate {□
        artifact ShipmentInformation {□
        artifact Products {□
        artifact SellerPayment {□
        artifact ShipperPayment {□
    }
}
```

Figure 5 Sample Process Model in ProcessDSL

```
interaction PlaceOrder {
    initiator Buyer
    responder Seller
    artifact RFI
    artifact Catalog
    artifact PO
    artifact POConfirmation
    artifact Payment
}
```

Figure 6 Sample Interaction Model in InteractionDSL

Regarding the tool for model transformation, the most widely used tools are eclipse/ATL and QVT. They could also be used to define and execute model transformations described in this paper.

The Xpand template used is shown in Figure 7. This should be considered as a sample. It does the following steps: a) import/read the metamodel, b) import the helper functions, c) declare this is a transformation against Process element, d) specify output file, e) define transformation processing for each lane, f) take only steps passing artifact to other lane, g) generate Service Interfaces, h) generate Participants, and i) generate Service Contracts with model information.

```
«IMPORT jp::ac::fun::xtext::process::processDsl»
«EXTENSION templates::Extensions»                                    b)
«REM» SoaML Service Architecture «ENDREM»                            c)
«DEFINE main FOR Process-»                                           d)
«FILE this.name+".soa"-»
ServiceArchitecture «this.name» {                                   e)
    «FOREACH this.lanes AS lane-»                                   f)
        «FOREACH lane.processFlows.typeSelect(Step) AS node-»
        «IF node.fromNode.fnode.eContainer()==node.eContainer()-»«ELSE-»
        «IF node.fromNode.isArtifactProvidedFrom()-»
    ServiceInterface «node.fromNode.fnode.name» {                   g)
    }
        «ENDIF-»«ENDIF-»«ENDFOREACH-»
    «ENDFOREACH-»
    «FOREACH this.lanes AS lane»«IF lane.name=="Artifacts"»«ELSE-»
    Participant «lane.name» {                                       h)
    }
    «ENDIF-»«ENDFOREACH-»
    «FOREACH this.lanes AS lane»«IF lane.name=="Artifacts"-»
    «FOREACH lane.processFlows AS msg-»
    Message «msg.name»
    «ENDFOREACH-»«ENDIF-»«ENDFOREACH-»
    «FOREACH this.lanes AS lane-»
        «FOREACH lane.processFlows.typeSelect(Step) AS node-»
        «IF node.toNode.tnode.eContainer()==node.eContainer()-»«ELSE-»
        «IF node.toNode.isArtifactProvided()-»
    ServiceContract «node.name» "" among {                          i)
        consumer «node.eContainer().getContainerName()»AsConsumer performed by
            «node.eContainer().getContainerName()»
        provider «node.toNode.tnode.eContainer().getContainerName()»AsProvider
            performed by «node.toNode.tnode.eContainer().getContainerName()»
        «node.eContainer().getContainerName()»AsConsumer-»«node.toNode.tnode.eContainer()
            .getContainerName()»AsProvider «node.toNode.anode.name»
    }
        «ENDIF-»«ENDIF-»«ENDFOREACH-»
    «ENDFOREACH-»
}
«ENDFILE-»
«ENDDEFINE»
```

Figure 7 Sample model transformation definition

As shown above, if the source model file contains enough information, and if the source model has defined grammar or metamodel, it is possible to create text file based on available information and by navigating the model elements. We have generated a small number of text files using this method, and will examine those in the next section.

One question people may ask is how we can make sure that the DSL targeting only core concepts can be used in the research like this. It is different, but if it contains major elements with right relationships, it is possible to compare the grammar or generated EMF ecore file with that of full modeling language to see if there is any major inconsistencies or issues in using it in the research.

9   RESULTS OF PROCESS TO SERVICE TRANSFORMATION

Now let us examine what the transformation produced. The first is a result around Service Architecture.

A.   Participant/Service Contact/Service Architecture

From the used template above, here is a summary of what operations were given to the sample process model.

1)   Service Architecture name is derived from Process's name.

2)   Participant name is derived from lane's name.

3)   Service Contract is defined for the step that passes Artefact across the lanes. Other types of actions will be

discarded. Service Contract name is derived from the step name that initiates interaction.

If we modify the grammar to include marking to show the logical boundary of the application, it may become possible to have coarse grained Service Contract with multiple Artefact, but that will introduce another requirement on the dependencies between Artefact.

The generated textual model is shown below (Figure 8), which is generated from sample process model in ProcessDSL (Figure 5) by applying process to service model transformation (Figure 7). Note that the model data below has been imported into ServiceDSL editor.

B. Service Interface

Service Interface in SoaML is functional elements, and is more like interface and signatures in Computational Object. The best way is to define Computational Object with process definition, but that will lead to a different world. Based on the generated Service Contract, it is safe to assume that Participant on receiving side have capability to process the Artefact passed from the other side. This implies that there exists Service Interface on the Service Provider Component. However there is no information about signatures in process models in general, it would not be possible to generate Message elements either.

4) Service Interface is derived from the node (name) that receives Artifact.

C. Service Choreography

Service Choreography is a set of defined sequence of service interactions between the two Participants, which is specified using UML Sequence Diagram. This contrasts with UML Activity Diagram we used to specify a process. Although not included in Figure 8, it is possible to collect interactions between different lanes in the process diagram.

Although BPMN, Activity Diagram, and SoaML are all graphical modeling language, we applied our method to define simplified textual DSL, and were able to transform a process model to a service model in a limited manner.

10 RESULTS OF SERVICE TO PROCESS TRANSFORMATION

In this section we will start with service model definition. The first step is to define the grammar for ServiceDSL that implements SoaML's core concepts, which are Service Architecture, Service Contract, Participant and Service Interface. Again, here is a portion of the grammar definition (more than ten elements are not shown in Figure 9).

This DSL is simple enough to cover the structural aspects of SoaML model, and we even tried to include behavioral (sequence diagram) aspect in the grammar.

Next thing is to create a service model based on this DSL, which is shown in Figure 10, which is done also by typing.

Figure 8 Sample Generated Service Architecture

Figure 9 Partial ServiceDSL grammar

The previous Figure 2 showed a graphical representation of a sample Service Architecture at very high level. In the textual ServiceDSL model of Figure 10, we included major elements under Service Architecture, because Service Architecture works as a root of the model in this language. Each usage of the typical language element is shown at least once, but not all the elements are shown by using the folding option to make the Figures smaller to fit in this paper. Also

note that message sequencing is specified in Service Contract, e.g. with "buyer -> seller RFI optional" and the following messages with directions in Figure 11.

The major difference between this service model description and the previous process model description is in the style of control flow description, i.e. sequence versus activity, and the number of parties involved, i.e. only two parties in service model vs. possibly more than three parties in process model. In service model, service is the central concept and therefore major players in service model are consumer and provider. On the other hand, in process model, the focus is on control flow, object flow, conditional or parallel split and merge, covering all the players that could be more than three players. It is clear that service model is not able to express e.g. control flow within the same lane in process model, since they are not exposed as service interaction and of no interest in service model.

```
ServiceArchitecture BuySellShip {
    ServiceInterface BuyInterface {
        RequiredInterface RFI {
            Parameter in RequiredProducts : String
            Parameter out Information : Document
        }
        RequiredInterface PO {
        RequiredInterface PaySeller {
    }
    ServiceInterface ReceiveProduct {
    ServiceInterface SellInterface {
    ServiceInterface ShipperInterface {
    Participant Buyer {
        RequestPoint RFI : BuyInterface
        RequestPoint PO : BuyInterface
        RequestPoint Payment : BuyInterface
        ServicePoint Products : ReceiveProduct
    }
    Participant Seller {
    Participant Shipper {
    ServiceChannel BuySell {
        Buyer.RFI -> Seller.RFI
        Buyer.PO -> Seller.PO
        Buyer.Payment -> Seller.ReceivePayment
    }
    ServiceChannel SellShip {
    ServiceChannel ShipBuy {
    Message RFI
    Message Catalog
    Message PO
    Message POConfirmation
    Message RequestShipping
    Message EstimateForShipping
    Message ShippingStatus
    Message DeliverProduct
    Message SellerPayment
    Message ShipperPayment
    ServiceContract PlaceOrder "Ordering Products" among {
        consumer buyer performed by Buyer
        provider seller performed by Seller
        buyer->seller RFI optional
        buyer<-seller Catalog
        buyer->seller PO
        buyer<-seller POConfirmation
        buyer->seller SellerPayment
    }
    ServiceContract RequestShipping "Seller request for shipment" among {
    ServiceContract DeliverProducts "Shipper delivers products" among {
}
```

Figure 10 Sample Service Architecture

We can still apply model transformation to see what we can get even though the limitation is clear. Figure 11 shows a transformed sample process based on the service model (Figure 10). It seems step portions of the process were successfully generated. But these are just concatenation of the sequences from Service Contract's sequence definitions.

If full control flow needs to be generated from the service description, process oriented description should be a part of the service model. In SoaML specification, these process aspects are treated as requirements specification to services, and therefore they are outside the scope of SoaML language itself (no stereotype is defined against Activity).

```
process BuySellShip {
    role Buyer {
        step RequestRFI
        step ReceiveCatalog
        step RequestPO
        step ReceivePOConfirmation
        step RequestSellerPayment
        step RequestShippingStatus
        step ReceiveDeliverProduct
    }
    role Seller {
        step ReceiveRFI
        step RequestCatalog
        step ReceivePO
        step RequestPOConfirmation
        step ReceiveSellerPayment
        step RequestRequestShipping
        step ReceiveEstimateForShipping
        step RequestShipperPayment
    }
    role Shipper {
        step ReceiveRequestShipping
        step RequestEstimateForShipping
        step ReceiveShipperPayment
        step ReceiveShippingStatus
        step RequestDeliverProduct
    }
    role Artifacts {
        artifact RFI
        artifact Catalog
        artifact PO
        artifact POConfirmation
        artifact RequestShipping
        artifact EstimateForShipping
        artifact ShippingStatus
        artifact DeliverProduct
        artifact SellerPayment
        artifact ShipperPayment
    }
}
```

Figure 11 Sample Transformed Process

11 RELATION WITH DISTRIBUTED COMPUTING

If we look at both process oriented and service oriented models in the context of distributed computing, they can be considered as candidate sources of execution on the platform, which may be internal enterprise systems or hybrid with cloud computing platform environment or full public cloud platform.

There is a category of technology called process engines that interpret and execute process definitions. Workflow engines are also considered as ones in this category. Their focus, however, is on controlling and monitoring the given process flow, and not on the execution of distributed services. Regarding SOA, Web Services is one of the typical implementation technologies, and they can be considered as base technology for distributed computing. With the use of SoaML, most of the necessary information to map down to SOA implementations is included in the model, and therefore it is not surprising to find products to do code generation based on SoaML model and actually run on the SOA runtime platform.

Our interest here is how close to implementation we can get based on process model via service model. We created a transformation template (not included in this paper) to generate skeleton interface codes of the service components. The result is shown in Figure 12. As expected, there is not much detailed information included, since some of the control information is discarded when converting it to service model. To make this code work, it needs to be completed with more detailed information, with

implementation classes, WSDL files, and frameworks for SOA such as eclipse SOA Platform etc.

```
package jp.ac.fun.xtext.services;

import javax.jws.*;

public interface SendCatalogService {
    public void SendCatalog (
    );
}
public interface AcceptPOService {
    public void AcceptPO (
    );
}
public interface RequestForInformationService {
    public void RequestForInformation (
    );
}
public interface SendPOService {
    public void SendPO (
    );
}
public interface SendEstimateService {
    public void SendEstimate (
    );
}
public interface PaymentToSellerService {
    public void PaymentToSeller (
    );
}
public interface ReqeustShippingService {
    public void ReqeustShipping (
    );
}
public interface PaymentToShipperService {
    public void PaymentToShipper (
    );
}
```

Figure 12 Generated Skeleton Code

12 CONCLUSIONS

DSLs are usually used at area close to programming. We demonstrated that DSLs, which captures only core concepts of the target modeling language, can be applied to architecture descriptions that are Process-oriented and Service-oriented architectures and can be used to examine the difference between the two styles of modeling presented in this paper.
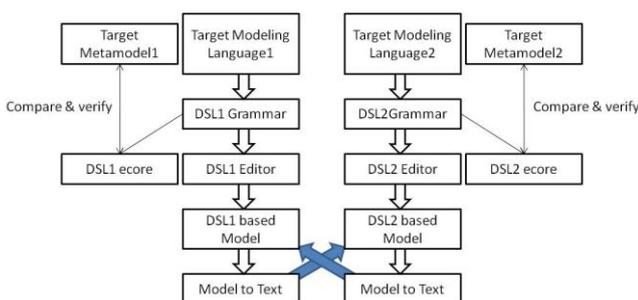
Figure 13 Model Transformation using M2T transformation

We created textual DSLs, including ProcessDSL and ServiceDSL, and showed sample transformations from process model to service model, and service model to process model. In doing so, we found a major difference between process modeling and service modeling. Something internal in process modeling will be lost when it is converted into service model, e.g. internal control flow. Orchestration of all the participants, which is the essential part of process model, is not possible in most cases when transforming service model to process model, since services are only meaningful to consumers and providers, i.e. between two parties, and normally orchestration aspects are left to higher level activities.

There is also a fundamental difference between the two, which is about level of abstraction. In process modeling, the level of abstraction is at end users or at business analysts level, but service modeling further includes interface specifications that are at architects' or developers' viewpoint. This caused transformation loss from service to process, and also was the reason of insufficient output from process model.

The possibility of service interface generation from process model was examined and only skeleton interface codes were generated because of the semantic gap between the two models with associated information loss in transformation. This does not, however, preclude the possibility of code generations from process model into process engines' environment and from service model into SOA environment.

Based on above, authors believe that they are showing the different aspects of the business system, and if mixed use is required, positioning process model as higher than service model will better work in enterprise architectures than positioning them in the opposite order or placing them at the same level.

Regarding the tooling, eclipse/Xtext provided necessary DSL development environment, and integrated model to text transformation facility Xpand/Xtend worked well to generate text from the DSL based models.

13 FUTURE WORKS

There are some areas where we need further works.

We will need to investigate a mechanism to verify the created DSLs, a mechanism to store trace/log information in transformation, more resources like practical examples around DSLs and model transformations, and more specialized tools to achieve specific activities.

We will also need to experiment on minimizing transformation loss such as internal steps in a process, possibly by introducing control flow manager in each lane so that each step could be transformed to internal service etc. or by introducing service orchestration function to the service model.

We are planning to use full RM-ODP model as source in the next experiment to see how component definitions in computational specification could contribute to process to service model transformations.

ACKNOWLEDGEMENTS

REFERENCES

[1] Zachman, J.A., A Framework for Information Systems Architecture. IBM Systems Journal, Vol 26 (3) 1987

[2] FEA Reference Models, Consolidated Reference Model Version 2.3, http://www.whitehouse.gov/sites/default/files/omb/assets/fea_docs/FEA_CRM_v23_Final_Oct_2007_Revised.pdf.

[3] OMG, Business Process Model And Notation (BPMN) Version 2.0 http://www.omg.org/spec/BPMN/2.0/PDF/

[4] OMG, Semantics Of Business Vocabulary And Business Rules (SBVR), Version 1.0, http://www.omg.org/spec/SBVR/1.0/PDF/

[5] Luckham, D: The Power of Events: An Introduction to Complex Event Processing in Distributed Enterprise Systems, Addison-Wesley, 2002

[6] OMG, Unified Modeling Language™ (UML®), http://www.omg.org/spec/UML/

[7] ISO/IEC 10746-2:2009, Information technology -- Open distributed processing -- Reference model: Foundations

ISO/IEC 10746-3:2009, Information technology -- Open distributed processing -- Reference model: Architecture

ISO/IEC 15414:2006, Information technology -- Open distributed processing -- Reference model -- Enterprise language

ISO/IEC 19793:2008, Information technology -- Open Distributed Processing -- Use of UML for ODP system specifications.

[8] Fowler M., Domain-Specific Languages, Addison-Wesley, 2011

[9] Kleppe A., Software Language Engineering, Addison-Wesley, 2009

[10] OASIS, Reference Model for Service Oriented Architecture 1.0, http://docs.oasis-open.org/soa-rm/v1.0/soa-rm.pdf.

[11] OMG, Service oriented architecture Modeling Language (SoaML®)

[12] OMG, MOF™ Query / View / Transformation (QVT), http://www.omg.org/spec/QVT/1.1/PDF/

[13] Jouault F., Allilaire F., Bezivin J., Kurtev I., ATL: A model transformation tool, Science of Computer Programming, Volume 72, Issues 1-2, Special Issue on Second issue of experimental software and toolkits (EST) (2008)

[14] Eclipse, ATL Transformations, http://www.eclipse.org/m2m/atl/atlTransformations/

[15] OMG, : XML Metadata Interchange (XMI®), http://www.omg.org/spec/XMI/

[16] Eclipse Foundation, Xtext - Language Development Framework, http://www.eclipse.org/Xtext/

[17] The Spoofax Language Workbench, http://strategoxt.org/Spoofax

[18] Meta Programming System, http://www.jetbrains.com/mps/

[19] Xtext Users Japan, http://sites.google.com/site/xtextusersjapan/files

Inspection Method of Privacy Utilization for Rule-based Mobile Applications

Chiaki Doi, Tomohiro Nakagawa, Ken Ohta, and Hiroshi Inamura

Research Laboratories, NTT DOCOMO, Japan
chiaki.doi.tf@nttdocomo.com

Abstract – As smartphones that offer the open development environment prevail, the number of applications is growing exponentially across the globe. Among them, context-aware mobile applications are increasing because various sensors are available to catch the user's context changes.

In order to create context-aware mobile applications, we adopt the ECA rule-based approach, which is intended for reactive applications. The application can perform functions without user's intervention, by leveraging context data such as terminal logs or sensor data as a trigger to perform predefined actions. Because context data can involve privacy data, we must ease privacy concerns before these applications will be widely accepted.

The objective of this research is to realize a method that determines how privacy data is utilized in context-aware rule-based mobile applications. The method outputs a privacy report, which is shown to users so that they can confirm if the application is acceptable or not in terms of privacy. There are three requirements that prescribe how privacy data utilization is to be described in the report.

The challenge of inspecting rule-based applications is to analyze the information flow on the implicit chain formed by independent rules. Such an implicit chain exists because there are cases in which the firing of one rule depends on another, even if the rule doesn't explicitly refer to the other.

Our privacy data inspection method is composed of a chain analysis process, a filtering process, and a summarization process. The proposed method satisfies the three requirements for privacy reporting, as well as accuracy and conciseness.

In order to evaluate the chain analysis and filtering processes, we create an evaluation model composed of two rules, which is the minimum rule chain unit. We confirm that neither false positives nor false-negatives are detected in evaluations that uses all 400 combinations of the rules generated from the model. We also confirm that the summarization process is effective in creating concise privacy reports. From this result, it is concluded that the proposed method works correctly for rule-based applications composed of straight chain rules.

Keywords: ECA rule, Information flow, Privacy data

1 INTRODUCTION

Due to the increasing number of mobile phone functions now open to application developers, it is becoming more attractive to create context-aware mobile applications. These applications leverage terminal logs such as user's location information or sensor data and behave autonomously by taking the user's context into account. For example,

BreadCrumbz [1] , Locale [2] , and Nakamap [3] use such information to navigate people to specified places

.

While the application market is expanding rapidly, concern is growing as to the security of using 3rd party applications. In recent studies of the Android platform, it has been pointed out that the Android framework faces high-risk threats[1][2], and that it is essential to determine how well applications can guarantee the Android user's security and privacy[3][9].

In order to create context-aware applications, we adopt the ECA rule, which is a suitable way to describe an application that automatically executes various functions depending on the user's situation.

Based on this ECA-rule based approach, we tackle the challenge of application inspection. Because the applications utilize the user's privacy data such as terminal log and sensor data, they suffer the risk of disclosing privacy information. For example, the application may directly pass the user's privacy data such as location information to another user. As another example, the application may indirectly convey the user's context to other users in the case that the user's context is utilized as the trigger to execute notification function to the other users.

In this paper, we propose a method to inspect how privacy information is utilized in ECA rule-based mobile applications. The proposed method is composed of a chain analysis process, a filtering process, and a summarization process. After these three processes conclude, the result, a privacy report, is shown to the user, The user can then confirm whether the application exhibits any harmful behavior or not before installing it.

In order to evaluate the chain analysis and the filtering processes, we created an evaluation model composed of two rules, which is the minimum rule chain unit. An evaluation result based on evaluations of all 400 combinations of the rules generated from the model shows that neither false positives nor false-negatives are detected. Thus it is confirmed that the proposed method works correctly for straight chain rules, and thus covers simple applications.

In addition, we evaluate the summarization process, and the result shows that the amount of privacy data is nearly halved in the privacy report while keeping the information needed by the user for decision making.

The paper is organized as follows. Section 2 describes the

---

[1] BreadCrumbz   http://www.bcrumbz.com/
[2] Locale   http://www.twofortyfouram.com/
[3] Nakamap   http://www.appbrain.com/app/nakamap-where-are-you-now/com.kayac.nakamap

ECA rule-based application and discusses related work; it focuses on existing static analysis techniques and tools. Section 3 describes three requirements for the privacy report to be accurate and concise. Section 4 introduces our proposed method. In this section we explain how to analyze the chain of the rules, how to filter unnecessary rule chains from the candidates, and how to summarize similar items to make the privacy report concise. Section 5 describes the evaluation model for straight rule chains, i.e. no branches or loops are present. Also, the result of an evaluation of the summarization method is shown in this section. Section 6 concludes with a summary and an outlook on future work.

2 ASSUMPTION AND RELATED WORK

2.1 ECA rule-based applications

In this study we adopt the ECA rule[4][5] to describe context-aware mobile applications, which autonomously execute functions depending on the user's context. The ECA rule is composed of an event, a condition, and an action. The event triggers the processing of the rule. After the event occurs, the specified condition is checked. Only when the condition is satisfied is the action implemented. The reason why we adopt the ECA rule is that the execution of the rule is conducted without the user's explicit intervention, which is a suitable attribute for context-aware applications.

A rule-based application is composed of a set of ECA rules, which are described in XML. Each ECA rule in the application works independently, and has a different trigger as defined by the event. Thus, the rule has atomicity in that its logic is self-contained. Due to rule atomicity, a ECA-rule based application has flexibility in terms of execution, which is an additional reason why we adopt the ECA rule. For example, previous research mentioned that an ECA-based application can be customized by adding or deleting rules [6]. Also, there is research on dynamic changing rules at runtime in a transparent manner [7]. We consider that part of the ECA rules will, in the future, be transferred between a terminal and a server to realize dynamic load balancing.

In order to describe various types of mobile applications, we define several tags for the event and the condition. These tags are utilized to catch the change in the user's context through the user's operation log or sensor data. For example, an *OCCUR* tag is defined to catch the generation of a specified terminal log. As another example, a *SUM* tag is defined to evaluate if total number of terminal log generation is more or less than a specified value.

Also, various kinds of terminal logs are available as the evaluation target for these tags. There are several kinds of logs such as screen light-up or application start-up, which indicate the user's operation of the mobile phone. In addition, there are other kinds of logs such as location information or pedometer data, which indicate the user's behavior. By combining the aforementioned tags and these terminal logs, you can describe different kinds of context-aware applications.

In addition, we define a user-defined event as a kind of log. The objective of the user defined-event is to gather the user's context from their pressing a button on a dialog.

Suppose the user-defined event *BEING_TIRED* is bound to a dialog button. When the button is pressed, the corresponding user-defined event is issued and recorded as user's operation log. Afterwards, the log of *BEING_TIRED* can be utilized in the same way as the normal kind of log from the event or condition by using tags such as *OCCUR* or *SUM*.

As an example of ECA-rule based applications, let's take an information distribution application which provides tourists with information depending on their location and pedometer data as captured by mobile phones. Figure 1 shows an excerpt of an ECA rule from the application. The rule defines that the event is fired when the user's location is within 1 km from Venice, which is located at latitude 45.434336 north and longitude 12.338784 east. Then, as defined in the condition, the pedometer data is checked to determine if it exceeds 5000 steps. If yes, as defined in the action, an implicit Intent is issued by using the Intent system of Android, and the browser accesses the specified URL of a web service which distributes tourist information. When the browser accesses the 3rd party's server, it transmits the user's location and pedometer data to obtain the user's context, which is utilized to personalize the delivered information.

This sample rule shows that the user's location and pedometer value is sent to the information provider. Thus, the user will want to know what kind of privacy data is being utilized and how the data is being conveyed to 3rd parties, before using the application.

```
<rule id="1">
  <event>
    <center lat="45.434336" lon="12.338784"
        kind="LOCATION_INFORMATION">
      <less_than>1000</less_than>
    </center>
  </event>
  <condition>
    <sum kind="PEDOMETER" more_than="5000">
    </sum>
  </condition>
  <action>
    <coordinate intentType="implicit"
      data="http://URL1?location=Venice&pedometer=
5000"
        action="android.intent.action.VIEW"/>
  </action>
</rule>
```

Figure 1   An example of ECA rule-based application

2.2 Rule chain by dependency between rules

As we have seen in the previous section, an ECA rule-based application is described as an aggregate of independent rules. In other words, no rule explicitly refers to any other rule.

In spite of rule independency, there is an implicit dependency between the rules in specific situations. Such a situation occurs when the execution of a specific rule's action is a necessary condition for firing another rule's event.

If there is such dependency between the rules, user's privacy data may indirectly flow out of the terminal through the continuous execution of these rules. In order to clarify an output path of privacy data, we need to analyze the dependency between rules.

There are two cases of such rule dependency, which is called 'rule chain' hereafter.

In the first case, the first rule's action has a *LOG START* tag and the second rule has any kind of event tag which utilizes the same kind of log specified in the first rule. The *LOG START* tag means to start recording a kind of log. In the second case, the second rule's event has an *OCCUR* tag, which means to catch the generation of a specified log type, the second rule is fired only after the first rule is executed.

The second case of the rule chain is that the first rule's action has a *BUTTON* tag, which issues a user-defined event, and the second rule's event has any kind of event tag which utilizes the same user-defined event specified in the first rule. If the second rule's event has an *OCCUR* tag that is defined to count the number of the same user-defined event specified in the first rule, there is a chain between those rules.

2.3 Related work

For the purpose of tracking how sensitive data is handled by an application, some studies attempt to track the data at runtime[8][9]. In [8], sensitive data such as passwords or credit card numbers are tracked by simulating the whole system. TaintDroid[9] realizes a light-weight system-wide tracking system for mobile phones.

While these tracking systems have the same purpose, analyzing the flow of sensitive data, they use dynamic analysis, which assumes that the target application can be run in an emulation environment or a real system.

Our approach is to use a static analysis approach, which can analyze the information flow without executing the target application. By using static analysis, the user can notice the malformed behavior of an application before installing it.

There are several techniques for the static analysis of information flow that annotate programs handling confidential information [10][11]. Our proposal avoids such annotation. All of the privacy data is predefined as confidential information, and the flow of the privacy data is exhaustively analyzed. Then, whether the privacy data should be reported or not is judged by how they are utilized.

There are some techniques that offer the static analysis of data or programs without annotation [12][13][14]. These techniques are intended to verify the query data[12][13] or sequential execution of the code, which follows an explicit control[14].

The difference between our proposed method and existing static analysis techniques is we analyze the implicit chain formed by independent rules. This study is intended for applications that consist of independent rules. However, as described in Section 2.2, the rules can form implicit chains. Therefore, it is necessary to consider the dependence described in Section 2.2. To analyze an implicit rule chain, it

is necessary to analyze the pathways of information by considering the semantic relationships between rules.

From the viewpoint of analyzing rule-based applications, a previous study described the verification of an application's convergence[15]. However, its goal was to verify whether the application's running time would converge or not, and so cannot be applied to our problem.

Existing related tools are Permission Checker Security [4] and S2Permission Checker [5] to detect applications that might be spyware applications in Android phones. The difference between these tools and our method is that those tools merely show the possibility of anomalous behavior at a rough level of granularity. For example, these tools do not show the destination of the leaked information. On the contrary, our proposed privacy report provides users with a more detailed assessment of the behavior of the application as described in Section 3.

3 PRIVACY UTILIZATION INSPECTION

3.1 Requirements

After the application is inspected, the result will be presented to the user as a privacy report as shown in Figure 2. The privacy report is meant to be shown before the user installs the application. The user can then judge whether or not it is acceptable to utilize the application by confirming how the user's privacy data is utilized.

In order to clarify the detailed behavior of the application, there are three requirements that the privacy report must satisfy. The motivation of detailing the privacy utilization as described below is that there are various terminal logs, and a subtle change in how they are utilized may change the mind of the user.

1. Show which kind of privacy data is utilized and to which destination the information is sent
2. Details the conditions under which the privacy data was created
3. Eliminate redundancy from the privacy report so that the user can understand it easily

---

[4] Permission Checker Security
http://www.appbrain.com/app/application-permission-checker/jp.ne.neko.freewing.PermissionChecker
[5] S2 Permission Checker
http://www.androidzoom.com/android_applications/tools/s2-permission-checker_luge.html
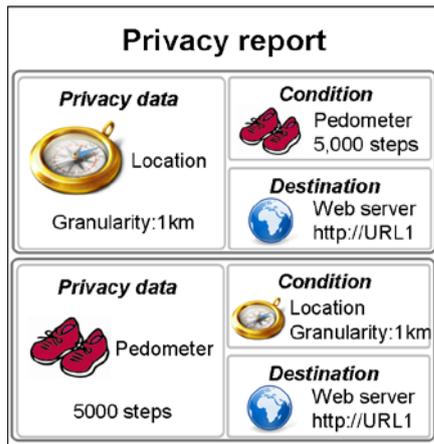
Figure 2 An example of the privacy report

Requirement 1 specifies that the privacy data such as user's location and pedometer data should be displayed jointly with the destination. The destination is specified for every notification method that uses the action. We defined eight kinds of actions such as e-mail transmission, Intent invocation, and location sharing. Existing tools specify only the privacy information that could be output. The addition of the destination makes the privacy report much more useful.

Requirement 2 specifies that the condition under which the action is taken should be included in the privacy report. We consider that the conditions may influence the user's judgment on whether or not to utilize the application. This is because the user should weigh the benefits and the privacy risks of the application in making the judgment. For example, suppose it is unacceptable for user A to share location information anytime and anywhere with person B. By limiting the time to share the location to around the time of rendezvous with B, the same application can turn to be an attractive one for the same user A.

Requirement 3 specifies that the privacy data described in the privacy report should be concise, eliminating items of the same sort. This is necessary to make the report readable.

3.2 Proposed inspection method

By adopting the ECA rule-based approach, we can describe a context-aware application flexibly by combining various kinds of tags and terminal logs. We designed an application programming language composed of seven events, three conditions, and eleven actions. In addition, 54 terminal logs can be utilized from the event and condition tags.

In order to generate accurate privacy reports, we need to clarify the implicit rule chain formed by any combination of the rules freely defined by utilizing these tags and logs. If there was any false recognition or oversight of a rule chain, the privacy report might contain false positives or false negatives.

In this section and the following subsections, we describe the proposed method, which is composed of the three processes shown in Figure 3. The proposed method processes the ECA rules and generates a privacy report as its output. The chain analysis process and second filtering

process are designed to satisfy requirements 1 and 2, by clarifying the sequence in which the privacy data is utilized. The summarization process is designed to satisfy requirement 3, by aggregating the same kind of items in the privacy report.
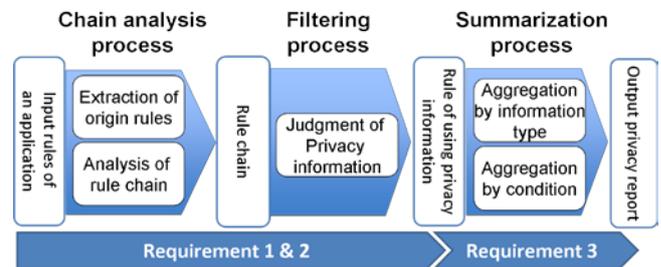
Figure 3   Processes of the proposed method

3.2.1.   Chain analysis process

The rule chain analyzed in this process starts from an origin rule and the chain is formed by the dependency of the rules as described in Section 2.2.

The origin rule is defined as a source of the privacy data utilization. In other words, it is the rule that utilizes the user's privacy data in its event, condition, or action tags. For example, a rule that includes *OCCUR* tag or a *SUM* tag is regarded as an origin rule. On the other hand, if a tag doesn't utilize any privacy data, a rule that uses only that tag is not an origin rule. For example, *TIME* tag is fired at the designated time regardless of the user's situation. That's why a rule that includes *TIME* tag is not an origin rule.

Note that the origin rule is defined by not only the event tag, but also the condition and action tags. The reasons are as follows. The condition tag has a function of judgment by taking the user's context data as a parameter, just as the event tag does. The action tag doesn't have such function, as it either directly or indirectly transmits the user's privacy data. An example of direct transmission is a *LOCATION* tag, which has the function of sending location information to a server to share it with designated agents/persons. An example of indirect transmission is a *BUTTON* tag, which has the function of issuing a user-defined event. In this case, the user's button operation can be transmitted via a rule chain.

Figure 4 shows a flowchart of the chain analysis process. This procedure is performed in two steps.

Category	Example of privacy data
User data	Phonebook entry added/modified
	Schedule added/deleted/modified
Terminal state	Battery power changed
	Silent mode set/canceled
Operation log	Screen light on/off
	Button operation
Behavioral information	User's location
	Pedometer data

Table 1 Example of privacy data in an origin rule

STEP 1
Extraction of origin rules

STEP 2
Analysis of rule chain
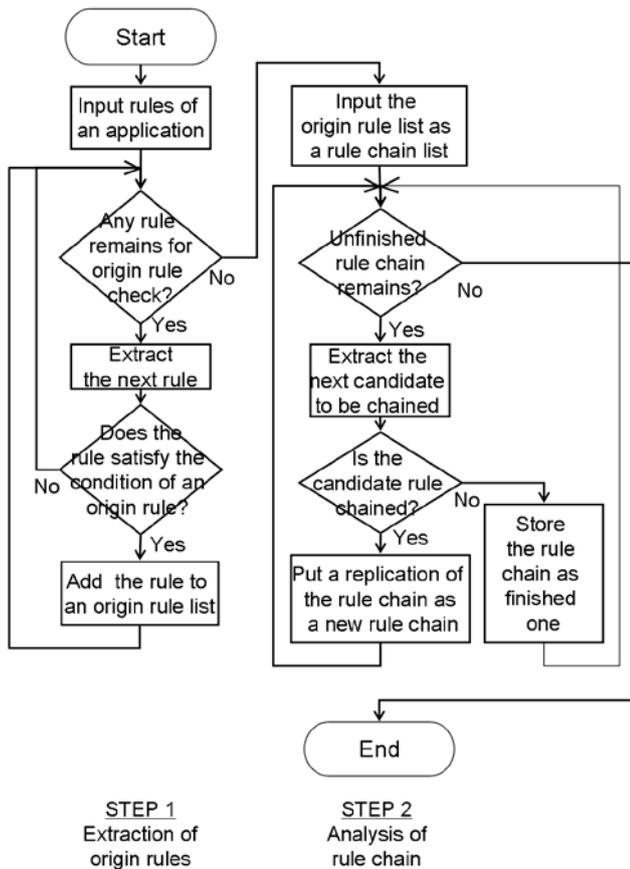
Figure 4 Flowchart of chain analysis process

The first step is to extract all origin rules, which is realized by pattern matching of the tags that utilize privacy data. There are four classes of privacy data: user data, terminal state, operation log, and behavioral information (see Table 1). These categories are utilized in the filtering process, as described in the next subsection.

The second step is to determine the rule chains starting from the origin rules extracted in the first step. This determines the pathway on which the user's privacy data is output. To extract all rule chains, whenever a new branch is found, the original rule chain is replicated. Loops are avoided by excluding rules that already exist in the rule chain.

The tricky part of the second step is that in some exceptional cases we need to define a rule chain as invalid, considering the semantics of the tags. For example, because *LOCATION* tag directly outputs user's location information by itself, a rule with the tag must be regarded as a finished rule chain, and so can not be connected to any other rule.

3.2.2. Filtering process

At the end of the rule chain analysis, we get a set of rule chains, which are filtered to extract the items essential for the privacy report. The motivation of the filtering process is to remove trivial items from the report. Even if a rule chain starts from privacy data utilization, the privacy data need not

be present in the privacy report as long as the data is utilized internally or in a normal way.

For filtering the privacy data, we focused on the relationship between privacy data in the origin rules and how the data is finally utilized in end rules. The end rule is defined as a rule that includes an action tag with an output function.

	Utilization purpose		
	Network output	Display output	Internal processing
User data	○	–	–
Terminal state	○	–	–
Operation log	○	○	–
Behavioral information	○	○	–

○ : Described in the privacy report
– : Omitted from the privacy report

Table 2 Criteria for privacy data filtering

The criteria for the filtering process are shown in Table 2. The intersection of the categories of privacy data and utilization purpose shows whether the privacy data should be described in the privacy report or not. The utilization purpose is categorized as network output, display output, and internal processing. In our filtering algorithm, all privacy data that is output via the network is reported to the user. On the other hand, only operation log and behavioral information are reported when these privacy data are output via the mobile's display. This is because user data and terminal data can be displayed for other applications, and these cases are regarded as relatively insignificant.

Just as in the case of the chain analysis process, we need to be careful of slight differences in the semantics of tags in the filtering process. We need to disable an end tag when the origin rule doesn't output user's context in a single rule. For example, a *BUTTON* tag is an origin tag which shows user's button operation. Even a rule including the *BUTTON* tag has also an action tag with output function, the button operation should be omitted from the privacy report. This is because the *BUTTON* tag is a part of an action tag . Only if the rule with the *BUTTON* tag is chained to other rules and the chained rule includes an output function, should the button operation be described in the privacy report.

3.2.3. Summarization process

In the summarization process, we eliminate the redundancy of the information presented to the user to satisfy requirement 3. This process aggregates similar rule chains by classifying them into the same category. These rule chains are aggregated when the same operation and behavior information are output by the same function.

We implemented two methods for the summarization process. Table 3 shows the items to be considered in the summarization process. The first method aggregates the rule chains by considering the conditions of privacy information. The second method does not consider these conditions.

	Origin rule		End rule
	Privacy data	Condition	
Method 1	○	○	○
Method 2	○	—	○

○：Considered , — ：Not Considered

Table 3　Items considered in the summarization process
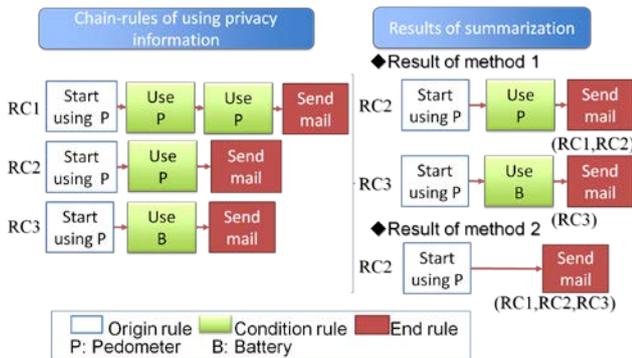
Figure 5　Operation of summarization process

Figure 5 shows the operation of the summarization process.

In method 1, judgment of whether the pedometer data meets the condition is performed for the second and the third rule. The second and third rules set the same condition; therefore, these rules are placed in the same category. RC1and RC2 are placed in one category if they have the same combination of privacy information and condition. As a result of summarization process, RC1 is aggregated into RC2 because RC2 has fewer rules than RC1. RC2 is output as the result of the summarization process. RC3 has different condition in the second rule, so RC3 is also output as an independent result.

In method 2, rule chains are aggregated without consideration of the conditions. Only the origin rule and the end rule are used for aggregation, so the amount of information output by summarization is reduced. The result of applying this filtering process to RC1, CR2 and RC3, is that RC1 and RC3 are aggregated into RC2 which has common origin rules and end rules.

The summarization process is stronger with method 2 than with method 1. The latter can show details that include the conditions of using privacy information.

4　EVALUATION

4.1　Filtering and Chain analysis processes

In order to evaluate the filtering process and the chain analysis process, we created an evaluation model composed of two rules, which is the minimum rule chain unit. Figure 6 shows the evaluation model proposed here. The rules in the evaluation model are composed of only event and action, i.e. they do not include any condition. The reason for this omission is that event and condition have the same mechanism for extracting privacy data.

For the sake of model completeness, the tags utilized in the evaluation model are selected so that they encompass all the categories that can impact the proposed processes. The result of the filtering process is influenced by the category; therefore we chose a representative tag from each category. We neglect the privacy data utilized in the tag because the difference in privacy data has no influence on the result of the chain analysis process.

Table 4 shows the tags used in this model. The event in the first rule includes four tags to cover all the categories of the privacy data which we defined in Table 1. The action in the first rule includes two tags which chain to other tags and three tags that output tags which include privacy information. The event in the third rule includes three tags which chained to other tags and two tags are chained to other tags. The action in the fourth rule includes three tags which output privacy information and a tag which finishes recording the log.
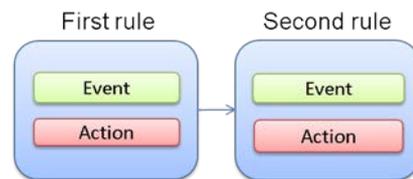
Figure 6 Evaluation model

Table 4　Tags in evaluation model

Rule		Function	Tags used in the evaluation model
The first rule	Event	Utilizes user's privacy data	<sum kind= "21">
			<sum kind= "43">
			<sum kind= "40">
			<sum kind= "28">
	Action	Chained to other tags	<log write>
			<log start>
		Output privacy information	<location>
			<dialog>
			<timer>
The second rule	Event	Chained to other tags	<occur kind="53">
			<occur kind="21">
		Not chained to other tags	<occur kind="43">
			<occur kind="40">
			<occur kind="28">
	Action	Not chained to other tags	<log stop>
		Output privacy information	<location>
			<dialog>
			<timer>

Our evaluation model effectively reduces the number of samples needed for the evaluation. Our model needs only 400 samples generated by the combinations of the representative tags. Without our model, 1,171,800 combinations would have to be evaluated.

An evaluation using all 400 combinations of the rules generated from the model confirms that neither false positives nor false negatives were detected. We covered

applications which chain linearly in this evaluation model. In other words, applications that include branches and loops are not covered. It is a future work to build a general evaluation model to cover such applications.

4.2 Summarization process

Our evaluation of the summarization process used the 18 kinds of tags shown in Table 4. We generated 25 applications, each of which had a set of five rules. The rule sets were generated by randomly choosing tag from among the candidate tag group with equal probability. The average number of the privacy data in the privacy report was 25 for the original data before the summarization process. After applying our summarization method, the number of the privacy data was reduced to 13 for method 1, and 11 for method 2. Thus, we confirmed that the summarization process is effective in reducing the number of private data, making it possible for users to confirm the privacy report more easily.

4.3 Processing time

We used Java to implement our proposed method. Table 5 shows the execution environment in which we evaluated our proposal.

Table 3 Execution environment

Execution environment	
CPU	IntelCore2 6400@2.13GHz
OS	Windows XP Professional SP3
RAM	2GB

Figure 7 shows the average processing time of the proposed method. The total processing time is the time required for filtering process, chain analysis process, and summarization process. The test data were the 25 applications utilized in the evaluation of the summarization process.

The average total processing time of the proposed method is 41 ms for method 1, and 37 ms for method 2. The result ranged from 32 ms to 44 ms for methods 1 and 2. This result confirms that either method 1 or 2 can be selected depending on the user's requirement.

	Method 1	Method 2
Summarization	4.5007	0.3518
Filtering	0.7008	0.7008
Chain analysis	35.5678	35.5678
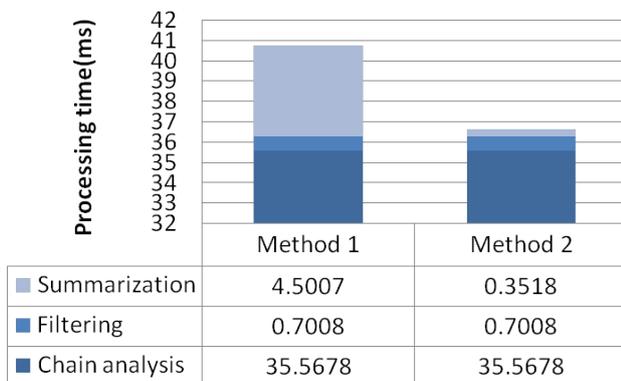
Figure 7 Processing time of the proposed method

5 CONCLUSION

We proposed an application inspection method for rule-based applications that is composed of a chain analysis process, a filtering process, and a summarization process. We confirmed the validity of the proposed method by using an evaluation model with minimum rule chain configuration. The evaluation results show that the proposed method works correctly for applications that are composed of straight rule chains. We also evaluated two summarization methods and confirmed that our method cut the amount of information present on the privacy report.

As a future work, we plan to evaluate the chain analysis and filtering processes by a more generalized model. In such a model, we need to include branches and loops of the rule chain to cover complicated applications. It is also a future work to improve the clarity of the privacy report in case there is a large number of privacy data.

REFERENCES

[1]A.Shabtai, Google Android: A Comprehensive Security Assessment, IEEE Security Privacy Magazine, pp.35-44 (2010).

[2]W.Enck, Understanding Android Security, Security Assessment, IEEE Security Privacy Magazine, pp.50-57 (2009).

[3]M.Butler, Android: Changing the Mobile Landscape, pervasive computing IEEE, pp.4-7 (2011).

[4]J.Bailey, An event-condition-action language for XML, Proceeding of the 11th international conference on World Wide Web, pp.486-495 (2002).

[5]G.Papamarkos, Event-condition-action rule languages for the semantic web, Workshop on Semantic Web and Databases,pp.855-864 (2003).

[6]M.Miyama, An event-driven navigation platform for wearable computing environments, IEEE International Symposium on Wearable Computers, pp.100-107 (2005).

[7]S.Chakravarthy, Dynamic Programming Environment for Active Rules, The 7th International Baltic Conference on Database and Information Systems, pp.3-16 (2006).

[8]J.Chow, Understanding Data Lifetime via Whole System Simulation. USENIX Security Symposium, pp.321-336 (2004).

[9] W.Enck, TaintDroid: An Information-Flow Tracking System for Realtime Privacy Monitoring on Smartphones, The 9th USENIX Symposium on Operating Systems Design and Implementation, pp.1-15 (2010).

[10]E.Dorothy, A lattice model of secure information flow, Communications of the ACM, Vol.19, No.5, pp.236-243 (1976).

[11]Andrew C. Myers, JFlow: Practical Mostly-Static Information Flow Control, The 26th ACM Symposium on Principles of Programming Languages, pp.228-241 (1999).

[12]P.Geneves, XML Reasoning Made Practical, The 26th International Conference on Data Engineering, pp.1169-1172 (2010).

[13]S.Boettecher, Finding the Leak: A Privacy Audit System for Sensitive XML Databases, The 22nd International Conference on Data Engineering Workshops, pp.100 (2006).

[14]C.Kirkegaard, Static Analysis of XML Transformation in Java, IEEE Transaction on Software Engineering, pp181-192 (2004).

[15]A.Cheng, Fast Static Analysis of Real-Time Rule-Based Systems to Verify Their Fixed Point Convergence, Proceedings of the 5th Annual Conference on Software Safety and Process Security,pp.46-56(1990).

Enhancement of Questionnaire on Anshin for Security
in the On-line Shopping by Web survey

Dai Nishioka[*], Yasuhiro Fujihara[*] and Yuko Murayama[*]

[*] Graduate School of Software and Information Science, Iwate Prefectural University, Japan
d.nishioka@comm.soft.iwate-pu.ac.jp,
fuji@iwate-pu.ac.jp,
murayama@iwate-pu.ac.jp

Abstract -We have conducted a user survey on Anshin with the users of information security technology and performed factor analysis. For the factor analysis, the factors with are deduced based on the question items in a questionnaire. Therefore, it is important to include items the question enough for factor analysis. We had surveyed Anshin factor on information security knowledge using questionnaire. Previous works used a questionnaire which was produced based on the preliminary survey with the computer science students. Since ordinary users using information security do not necessarily have technical knowledge, we needed a questionnaire to reflect the users without technical knowledge. We collected feedback from those without technical knowledge and produced a new questionnaire. We propose the use of brainstorming and the KJ method to review the questionnaire. Then, we produced a new questionnaire based on the results from brainstorming the KJ method. We conducted a pretest and with eighty one subjects using the new questionnaire. This paper reports our trial to produce a new questionnaire and the result from the pretest. The result shows the possibility of additional factors.

Keywords: Anshin, Factor analysis, Brainstorming, KJ method, Trust

1 INTRODUCTION

Anshin is a Japanese term that indicates the sense of security and safety. The concept of trust is investigated as the research which is similar to Anshin in Europe and USA [1] [2] [3]. Trust has been researched sociology, psychology and economics. Deutsch introduces to expect from the other person as confidence in the trust [4] [5]. Marsh proposed the computational trust model with trust values of -1 to +1[6]. We have identified that Anshin is the emotional part of trust [7]. We are currently working on Anshin to derive the factors of Anshin [8] [9]. We conducted the user survey with a questionnaire and perform a factor analysis on the survey responses. The original questionnaire was produced based on the responses from free description on a questionnaire for the students from Department of Software and Information Science in our university.

Ordinary users using information security do not necessarily have technical knowledge. When we conduct a survey on Anshin about information security for the ordinary people, it is important for a questionnaire to reflect the ideas of users without technical knowledge. For the factor analysis, the factors when a question items are insufficient are not extracted. We must collect the idea of the users without technical knowledge, and when the collected ideas arrange the question items, we must check to reflect the ideas with the users. However the users without technical knowledge do not understand the ideas of information security and knowledge of creation the questionnaire. Therefore we suggest to collect ideas from the users without technical knowledge and to create the questionnaire with the users.

In this research, we try and improve the questionnaire so as to obtain comprehensive coverage of question items to deduce Anshin factors. In this research, we incorporate brainstorming [10] and KJ method [11]. The brainstorming is used for generating ideas by a group of people. KJ method is an information organization method. Matching words as well as phrases in this process is based on semantics non letters. The next section presents our previous work. Section 3 reports methods to produce new question items on Anshin. Section 4 reports to create new questionnaire. Section 5 reports our recent web survey on Anshin . Section 6 reports discussion. Later sections describe the results of experimental survey including factor analysis. The final section gives some conclusions and future works.

2 PREVIOUS WORKS

Our previous works included user surveys on Anshin in terms of information security. The First survey, we conducted a questionnaire survey on Anshin when they use a security system or service on the Internet with four hundred and fifty two students of Iwate Prefectural University. Students are from various and divided into computer science students and non computer science students. As the result of the analysis, we found six factors: Security technology, Usability, Experience, Preference, Understanding and Cognitive Trust. Security Technology is based on the ability of a service provider from the users' viewpoint. Usability is the users' satisfaction of the user interface such as usability, attractive design and user-friendliness. Experience is concerned with how much the users have used the system and with recommendations of family and friends. Preference is concerned with the user's tastes such as interface design. Understanding is the users' knowledge about information technology or risk. Cognitive Trust is concerned with the users' confidence in company. Moreover, the former depend heavily on the knowledge for technology of a security system or service on the Internet,

the latter depend heavily on cognitive trust for service provider and preference for interface. However, about the users without information security knowledge, the number of data was not enough to analyze Anshin factors.

Later, we conducted another survey with users without the knowledge. However, original questionnaire was hard to understand the situation. From this, later survey was improved in the question items which were easy to assume a concrete situation. We set a situation of using online shopping and improved the questionnaire.

Result of the analysis, we found five factors: Cognitive Trust, Kindness, Understanding, Preference, and Familiarity. Cognitive Trust is concerned with the users' confidence in company.

With those surveys, we use a questionnaire which was produced based on the preliminary survey with the computer science students. Since ordinary users using information security do not necessarily have technical knowledge, we needed a questionnaire to reflect the users without technical knowledge. We collected feedback from those without technical knowledge and produced a new questionnaire.

3 METHODS TO PRODUCE QUESTION ITEMS

3.1 Overview

In our first survey [8], we conducted a preliminary survey for users with technical knowledge about information security technology and produced the questionnaire.

Ordinary users using information security do not necessarily have technical knowledge. When we conduct a survey on Anshin about information security for the ordinary people, it is important for a questionnaire to reflect the ideas of users without technical knowledge. However the users without technical knowledge may not understand the ideas of information security. We collected viewpoints from users without technical knowledge by conducting brainstorming with two groups of several people. Brainstorming is a technique to discuss with group and to collect the ideas. We thought that the users without technical knowledge can image the security knowledge using brainstorming and we can correct various ideas. The collected ideas using brainstorming is subjective ideas from user without technical knowledge. When the collected ideas arrange the question items, we must check to reflect the ideas of user without technical knowledge. However users without technical knowledge do not understand the question items to reflect the ideas of user without technical knowledge when we report the question items using technical knowledge. We sorted out the ideas by the KJ method. KJ method is a technique to arrange subjective ideas and to show the ideas as a figure. We thought that we can arrange subjective ideas from user without technical knowledge to create the question items and we can check the question items with user without technical knowledge using figure.

In this section, we report our first trial to produce a questionnaire. Our research process has five stages as shown in Fig 1. This process has been created along the process to create a general questionnaire. The process to create a general questionnaire has three processes [12]. The first process is to collect the ideas, the second process is to create the questionnaire, the third process is correct the questionnaire. Our process correspond the first process and the second process.

At the first stage, we collect ideas of feeling Anshin which the user inputs personal information on online shopping by brainstorming. Second stage, we arrange the ideas by KJ method. Third stage, we create the question items. The ideas are not a form of the question items on the second stage. Therefore we changed the wording of the ideas to be usable as a question items. Fourth stage, we compare the previous question items with arranging items and we collect the new question items which there do not include in the previous question items. Fifth stage, we create the new questionnaire that include new question items.
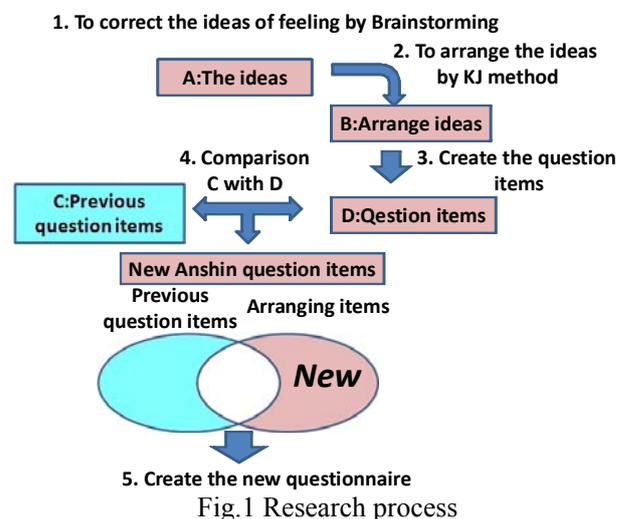
Fig.1 Research process

3.2 The Brain storming and the KJ method

Brainstorming is used for generating ideas with a group of people rather than with an individual. Brainstorming has four rules as follows:
1. Focus on quantity
2. Without criticism
3. Welcome unusual ideas
4. Combine and improve ideas

As far as brainstorming is concerned the quantity is more important than the quality. The participants of the brainstorming are required to present as many ideas as possible. The second rule is that the participants do not allow denying the others' ideas. The third rule is that unusual ideas should be welcome. The fourth rule is that ideas are combined to form a single better idea.

The KJ method is a tool used to organize ideas and data. It was created in the 1960's by Jiro Kawakita in Japan. The KJ method is commonly used within project management and allows large numbers of ideas to be sorted into groups for review and analysis, KJ method process has four steps as follows:

1.   Card making
2.   Grouping and naming
3.   Chart making
4.   Explanation

Card making is to write ideas on 4 X 6 cards. At grouping and naming, we shuffle the cards, spread them out and read them carefully. Cards that look as though they belong together semantically should be grouped. For each group write a title and place it on top of its group of cards. If we have more than about 10 groups, repeat these grouping and naming processes again and again. We assign a symbol mark to the final label. A symbol mark could be a sentence or a figure to indicate intuitively what the group is. At chart making, we make chart that contains the relations between groups on a large sheet of paper. At the Explanation, We make a presentation and a research paper to explain the meaning of the figure which we made. The KJ method could be used with the other idea creation methods such as idea creation method, idea creation: brain storming and mind map [12].

4   NEW QUESTIONNAIRE

4.1 Create the New Questionnaire

We conducted the brainstorming to collect users' images of Anshin. This work is first stage. We asked the subjects what makes them feel Anshin when they use the credit card at on-line shopping. We conducted experiment twice with different subjects. First group survey was conducted on five subjects of using on-line shopping with the credit card and without information security knowledge, in 25 and 27, November, 2008. Second group survey was conducted on five subjects of using on-line shopping with the credit card and without information security knowledge, in 24 and 25, January, 2009. Second group subjects were five office administrators of Iwate Prefectural University. We asked for their ideas about Anshin when using online-shopping. Each experiment took eight hours in total for 4 hours on 1 day. As a result of the brainstorming, we had in total 71 ideas. First group ideas were 32 ideas. Second group ideas were 39 ideas.

However, these ideas were some same ideas and some ideas were difficult to understand sentence. Therefore, we arranged 71 ideas by Card making on KJ method. This work is second stage. At first we revised the ideas that were difficult to understand a sentence and we wrote revising ideas on 4 X 6 cards. Second, we organized the same cards semantically by Grouping and naming on KJ method. We performed Grouping and naming once for compare this work with the previous question items. As a result, we had 46 items. We arranged 46 items for creating the question items to compare previous question items. 46 items is the feedbacks. These items must arrange question items. This stage is forth stage. We extract the new question items which were not included in the previous questionnaire. This work is fourth stage. The previous questionnaire was 28 question items. We compared these items with 46 items. As a result, 24 items were not included in the previous questionnaire.

We created a new questionnaire. This work is fifth stage. The new questionnaire would consist of 52 items. Previous question items ware 28 items. New question items were 24 items. However, if questionnaire is composed of many question items, subjects may become strained. In this case, subjects may not reply answer carefully. Therefore, we reduced question items to 36 items by Group and naming on KJ method. Previous question items ware 24 items. New question items ware 12 items. In the new questionnaire, we use a 7-point Liker scale, ranging from strongly agree (1) to strongly disagree (7). In addition, our some question items were same matter. Therefore we collected question items of the same matter and we created 6 groups. Group 1 consists of eight question items. Group 1 asks about the company. Group 2 consists of nine question items. Group 2 asks about the subjects. Group 3 consists of nine question items. Group 3 asks about the system. Group 4 consists of five question items. Group 4 asks about the on-line shopping. Group 5 consists of three question items. Group 5 asks about the service. Group 6 consists of two question items. Group 6 asks about the others. We create a new questionnaire as shown in Appendix A.

4.2   The result of grouping

We performed a continuance of the KJ method to investigate what kind of meaning the 24 items which we introduced into a new questionnaire. We organized the same cards semantically by Grouping and naming on KJ method. For on-line shopping, we have three players, viz, the user, the company and the information technology. We organized 46 items into the following seven groups by KJ method and we show relations of three players and seven groups as follows as shown in Fig 2:

1.   Advice of another person
2.   Information technology understanding
3.   Preference
4.   Risk understanding and measures
5.   Third Party
6.   Cognitive trust to the company
7.   Knowledge except information technology

Advice of another person is concerned with recommendations such as a testimony by the third person, a good experience by the family or the friend, word-of-mouth, the reputation system on the Internet. Information technology understanding is the users' knowledge about information technology such as good understanding of ICT, good understanding of security mechanisms, virus measure. Preference is concerned with the user's tastes such as preferred user interface, impressed design of web pages, and previous use of the site. Risk understanding and measures is the users' knowledge of risk and measures such as up-to-date technologies, competitive awareness of net crimes and defense against them. Third Party is concerned with third-party intervention such as compensation benefits by credit card companies. Cognitive Trust is concerned with the users' confidence in company such as direct management by trusted merchants, reputable/well established merchants.

Knowledge except information technology is the users' knowledge of on-line shopping such as company profile (e.g. location, capitals, executive officers, history), product information presented on the web site, prompt announcement of product recalls.

Groups 1 2 3 4 and 6 are related to the user, the information technology and the company, Groups 5 and 7 are concerned with the on-line shopping. We compared 24 items which were not included in our previous questionnaire. These items fit into two groups, viz, "Third Party" and "Knowledge except information technology". We introduce question items related to these two groups into our new questionnaire.
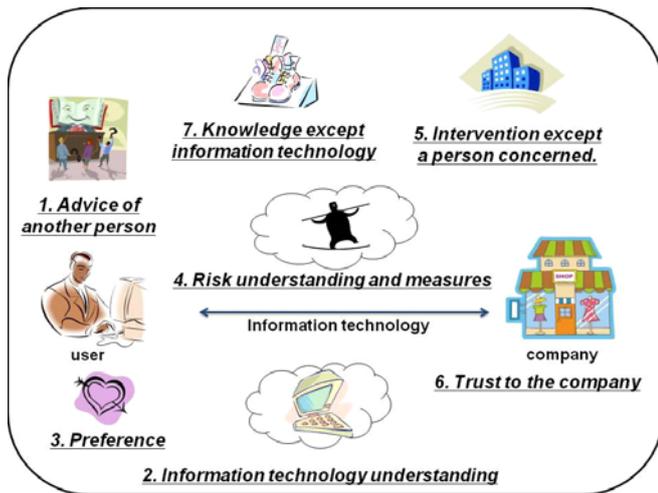
Fig.2 Seven groups by KJ method

5 DEADLINES AND AUTHOR KIT

We conducted a pretest using the newly created questionnaire through a web survey. The survey was conducted on 103 subjects of using on-line shopping with the credit card and without information security knowledge, in 22 and 23, July, 2010. We asked for their ideas about Anshin when using online-shopping. Of the 103 subjects, 39 out of 103 subjects were male, and 64 were female.

We analyzed the survey results using factor analysis. As the result of factor analysis with the maximum-likelihood method and the promax rotation found that five factors are derived as Tables 1. We resulted in the following factors:

1. Cognitive trust
2. Relation with the real world
3. Preference
4. Social position of the company
5. Understanding

Cognitive trust consists of 12 items which are concerned with the users' confidence in service providers and systems.

Additional Information through the other media consists of 10 items which are concerned with real world in service providers and systems.

Preference consists of 6 items which show user preference for interface design including one's tastes.

Social position of the company consists of 3 items, which are concerned with the users' confidence in the scale of the company

Understanding consists of 5 items which address the users'

knowledge about information technology.

Particularly, it shows their perception and understanding of risk or threat based on their prior knowledge. The above five factors include 36 items in total. The five factors were explained by 60.82% (Cumulative) of the total. To confirm reliability of measurement, Cronbach's coefficient alpha of each subscale Factor 1 was 0.922, Factor 2 was 0.793, Factor 3 was 0.856, Factor 4 was 0.927 and factor 5 was 0.800. Cumulative means how much all factors can explain all question items. Cronbach's coefficient alpha means stream that the users answer the same rating in a certain group such as all question items or the factors' items. If these are low: Cumulative is under 60%, Cronbach's coefficient alpha is under 0.80, we must consider question items or number of factors again.

As a result, Factor 1, Factor 3 and Factor 5 are similar to our previous survey results. Factor 1 was Cognitive Trust. Factor 3 was Kindness, Preference and Familiarity. Factor 5 was Understanding. Factor 2 is induced from many new question items and the factor loading was high. However, Factor 3s' Cronbach's coefficient alpha was low (0.793). Factor 4 consist of part of question group about the company.

Table.1 Factor pattern matrix

Factor name	Item No.	1	2	3	4	5
Cognitive trust	A2-5	0.868	0.051	-0.128	0.012	-0.082
	A2-3	0.847	-0.171	-0.305	0.086	0.139
	A3-2	0.815	-0.11	-0.022	-0.016	0.08
Additional Information through the other media	A1-6	-0.005	0.844	-0.017	-0.088	-0.073
	A1-8	-0.032	0.81	-0.046	-0.01	0.031
	A4-2	-0.133	0.71	-0.111	0.228	0.081
Preference	A3-8	-0.033	-0.062	0.933	0.091	-0.021
	A3-9	-0.041	-0.143	0.843	-0.065	0.063
	A3-7	0.041	0.15	0.704	0.127	-0.006
Social position of the company	A1-2	-0.107	0.139	0.059	0.957	-0.064
	A1-1	0.012	-0.098	0.071	0.915	0.045
	A1-3	0.085	0.129	-0.004	0.827	-0.12
Understanding	A5-2	0.019	0.113	-0.092	-0.002	0.827
	A5-3	-0.004	0.008	0.142	-0.086	0.761
	A5-1	0.062	0.066	0.007	-0.02	0.738

6 DISCUSSION

We extracted a factor which was not extracted at the previous survey. We examine whether this factor is a new factor of Anshin or not. This factor is "Additional Information through the other media". It consists of ten items, which is the additional information related to the service provider and the shopping item through the other media. However, this factors' Cronbach's coefficient alpha is so low as around 0.793 and the factor loading of the top four items are high, whereas the factor loading of other items are low in this factor. Cronbach's α is the internal consistency or reliability of a questionnaire survey score for

the subjects. Factor loadings of factor 2 are 0.844, 0.810, 0.710, 0.671, 0.395, 0.380, 0.361, 0.270, 0.266 and 0.219. The items with low factor loadings are low commonality. Commonality is the ratio to show how much an extracted factor could be explained by a question item. When a commonality is low, this question item becomes independent. Therefore, we look into the question items.

We produced the new questionnaire for the users without security knowledge. However, we transformed the feedbacks from the users without knowledge of the information security into the question items by ourselves --- i.e. the people with the security knowledge at third stage as shown Fig 1. That is, the new question items may well not reflect exactly the feedbacks from the users without knowledge, but had a bias from us. As the result of web survey, for some of the new question items we had the answers with the average value over 4. Moreover, for these new question items we had many answers with "neither disagree nor agree." Accordingly there is something wrong with the new question itmes and we need to review them. Therefore we need to look back to the original feedbacks of our KJ method process and reproduce the new questionnaire again examining them with the very subjects who gave the feedbacks so that the new question items will reflect the intention of the subjects without the security knowledge.

7 CONDCLUSION AND FUTURE WORKS

We used two techniques of brainstorming and KJ method to collect the new question items which there do not include in the previous question items for factors of Anshin. We extracted feedbacks of users' feel Anshin by the brainstorming and the KJ method. As a result, we extracted 46 items. We extracted the new question items which were not included in the previous questionnaire. As a result, we extracted 24 items. We created the new questionnaire to introduce 24 items. New questionnaire consisted of 36 question items. In this survey, we reflected on the ideas of the users without technical knowledge on the information security and the questionnaire. Similarly, it is thought that we can create the questionnaire using this technique when we conduct something survey on the users without technical knowledge.

We extracted seven factors which enters personal information surely about the online shopping by KJ method. Groups 1 2 3 4 and 6 are related to the user, the information technology and the company, Groups 5 and 7 are concerned with the on-line shopping. Groups 5 and 7 did not introduce previous question items. We conducted pretest using a questionnaire by web survey. As a result, we extracted five factors. However, new question items were problems. In the future works, we return to a result of the KJ method and must examine the question items which we introduced, we correct question items and investigate using the correcting questionnaire by Web survey.

REFERENCES

[1]Xiao, S. and Benbasat, I., *Understanding Customer Trust in Agent-Mediated Electronic Commerce*, Web-Mediated Electronic Commerce, and Traditional Commerce, Information Technology and Management, Vol.4, No.1– 2, Kluwer Academic Publishers(2004), 181–207.

[2]Xiao, S. and Benbasat, I., *The formation oftrust and distrust in recommendation agents in repeated interactions: a process-tracing analysis*, Proc. of the 5th international conference on Electronic commerce (2003) , 287–293.

[3]Gambetta, D., *Can we trust trust?,* Making and Breaking Cooperative Relations, electronic edition, Department of Sociology, University of Oxford, chapter 13 (1988) , 213-237.

[4]Deutsch, M., *Trust and Suspicion*, The Journal of Conflict Resolution, Vol. 2, No. 4 (1958) , 265–279.

[5]Deutsch, M., *The effect of motivational orientation upon trust and suspicion*, Human Relations (1960), 123–139.

[6]Marsh, S.P., *Formalising trust as computational concept*, Ph.D Thesis, Department of Mathematics and Computer Science, University of Stirling (1994).

[7]Murayama, Y., Hikage, N., Fijiwara Y. and Hauser, C., *The structure of the sense of security,Anshin,* Proc. of CIRITS2007 (2007), 85-96

[8]Hikage, N., Hauser, C. and Murayama, Y., *A Statistical Discussion of the Sense of Security, Anshin,* Information Processing Society of Japan Journal Vol.48 No.9 (2007), 3193-3203.

[9]Fujihara. Y., Yamaguchi. K., Y., Murayama. Y., *A Survey on Anshin of the Users without Technical Knowledge on Information Security,* Information Processing Society of Japan Journal Vol.50 No.9 (2009), 2207-2217.

[10]Osborn, A.F., *YOUR CREATIVE POWER*, Motorola *University Press*, abridged edition （1948）

[11]Kawakita, J., *The KJ method - A scientific approach to problem solving,* Technical report. Tokyo: Kawakita Research Institute (1975).

[12] Converse , J., Presser, S.: *Survey Questions: Handcrafting the Standardized Questionnaire*. Sage: Thousand Oaks, CA.(1986).

Session 3:
Network Routing and Multimedia
(Chair: Hideyuki Takahashi)

An Effective Lookup Strategy for Recursive and Iterative Lookup

on Hierarchical DHT

Tomonori Funahashi[*], Yoshitaka Nakamura[**] , Yoh Shiraishi[**] , and Osamu Takahashi[**]

[*] Graduate School of Systems Information Science, Future University Hakodate, Japan
[**] School of Systems Information Science, Future University Hakodate, Japan
{g2111032, y-nakamr, siraisi, osamu}@fun.ac.jp

Abstract - Recursive and iterative lookups on the performance of distributed hash table (DHT) are deteriorated by churn that nodes leave the network. When churn occurs infrequently, recursive lookup outperforms iterative lookup, but it returns back when churn occurs frequently. Therefore, optimal lookup needs recursive and iterative lookups to be separated by the frequency of churn. We propose a lookup strategy that separates recursive and iterative lookups by the churn rate. However, a common DHT makes it difficult t establish the neighboring churn rate. Hierarchical DHT takes into consideration the reliability of nodes to ascertain the churn rate. Therefore, we compared our lookup strategy with the use of either recursive or iterative lookup on hierarchical DHT.

Keywords: Recursive lookup, Iterative lookup, Hierarchical DHT

1 INTRODUCTION

Peer-to-Peer (P2P) is communication in which each node is equal and various values are dispersed throughout the network. Therefore, distributed hash table (DHT) is an efficient lookup technology in P2P. DHT can discover values with low numbers of hops in large networks. Examples of DHT based P2P include Chord [1], Kademlia [2], and Pastry [3]. Even if DHT uses the same algorithm as Chord or has routes on the same lookup path, their communication methods are defined differently. Its methods are known to be recursive and iterative lookups [4]. These lookups have different lookup latencies and numbers of messages. Recursive lookup, which has low latency, is generally satisfactory. However, the performance of these lookups deteriorates due to churn where nodes leave the network. In addition, recursive lookup performs worse than iterative lookup. Therefore, optimal lookup needs recursive and iterative lookups to be separated by the system churn rate. However, flat normal DHT it is not structured to take into consideration the feature of nodes, e.g. the churn of nodes. For this reason, it is difficult to establish the system churn rate.

There is a structure called hierarchical DHT [8][9] that enables DHT to be used efficiently. This structure can separate a number of clusters depending on needs. There is hierarchical DHT with advanced features that has taken into consideration how reliability of node is [10]. This has a clustering method that establishes the reliability of nodes.

Thus, each cluster is established the reliability approximately.

We propose applying an optimal lookup strategy to each cluster on hierarchical DHT that takes into consideration the reliability of nodes and separates recursive and iterative lookups efficiently in this paper.

2 RELATED WORK

2.1 Chord

Chord is a DHT algorithm that takes into consideration the hash space as a space like a ring, and sets nodes an identifier called the node ID with the hash function. Keys are calculated similarly with this function. Of the nodes arriving in a network, the node just behind a node is called a successor node, and the one just before a node is called a predecessor node. Nodes keep the neighbor as successor list which has a number of successor nodes, and a finger table that can route efficiently to the routing table. Chord completes path length $O(\log N)$ with these routing tables when N is the number of nodes. The state of these nodes is the previous state obtained by churn and failure. For this reason, Chord is implemented as a stabilization process to accurately retain the state of neighbor nodes. This is a process where nodes ask nodes in the routing table. In addition, it is executed at regular intervals.

2.2 Lookup strategy

Recursive lookup is a lookup strategy that originator node which demands value requests lookup other nodes. However, iterative lookup is a method which the originator controls lookup to ask other nodes about candidates for the next hop. Figure 1 outlines the shape of each lookup on Chord when the lookup has three hops (path length).

The originator in recursive lookup forwards a request message to a node that is closer to the destination (Figure 1 (1)). If a node received a request message does not have the purposed value, it forwards the request message to a node that is closer to the destination than itself. This process is executed till the request message reaches the destination node (Figure 1 (2), (3)). In contrast, the originator receives reply messages for request messages after the message has been forwarded in iterative lookup (Figure 1 (1-2), (2-2)).
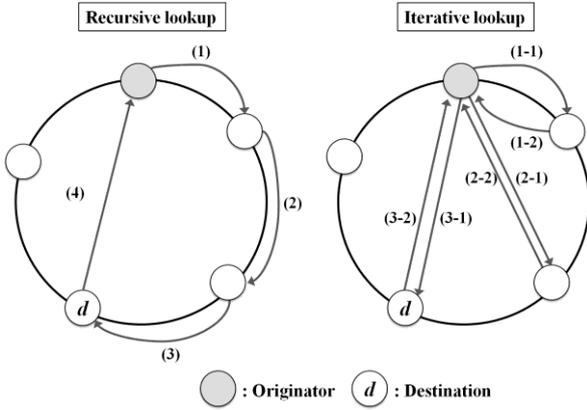
Figure 1: Recursive and iterative lookup strategy on Chord when path length = 3.

When a node received request message does not have the purposed value, it makes the reply message including the addresses of nodes which are closer to the destination than itself. The originator node forwards a request message to the destination node by using the address included in the reply message.

The performances of recursive and iterative lookups are affected by these communication methods and churn where nodes leave the network. The system churn rate, which is the probability what nodes will leave the network, is determined the life-time of nodes. R is the defined life-time of a node and refers to the reliability of nodes. R varies between nodes. The cumulative distribution function [5] of exponential or Pareto distribution [6] is used as a function to define R. R shows how often churn occurs in the system. S is defined as the time until nodes detect failure and repair the routing table of the node when churn or failure occurs. For this reason, S just means the interval in which the stabilization process is executed. Altogether, large S means that the stabilization process is seldom executed, but small S means that stabilization is executed often.

We also assumed that $E[R]$ and $E[S]$ were value expected for the R and S of neighbor nodes for a node. By using these parameters, p is defined as the probability of which next hop candidate node is alive in the network and the success of forwarding a request message, which is given by the following [7].

$$p = \frac{E[R]}{E[R] + E[S]} \qquad (1)$$

When neighbor nodes are in a steady state when starting lookup and the originator is not executed to repair its own routing table, $E[S]$ approximates a fixed value. As a result, p depends on $E[R]$. In addition, large $E[R]$ means that neighbor nodes are alive for a long time, and this also means that churn is not likely to occur. In contrast, small $E[R]$ means that churn often occur in neighbor nodes that have shorter lifetimes. That is, the churn rate is low when p is high, and it is high when p is low. More specifically, p means the churn rate in the network when $E[S]$ approximates a fixed value.

The performance of recursive and iterative lookups are defined by using churn rate p and latency of communication [7]. First, we assume that the lookup path length is l and t is the latency for one hop. We also assume that physical links between nodes are not considered, and t is fixed. In addition, T is the time, which is timeout when nodes fail to forward messages by churn or failure. Here, timeout T is configured differently at each lookup. The originator in recursive lookup has to wait for responses to complete as lookup is completed. However, other nodes only forward request message to the next hop node, and are not concerned with the forwarded message. Therefore, T in recursive lookup is set to no less than the time to complete the entire lookup at only the originator. For this reason, T_r as the timeout in recursive lookup is configured as $T_r \geq (l+1)t$. The originator in iterative lookup similarly waits for a response from the next hop node point by point. Therefore, timeout is configured to no less than the time to wait for forward and reply. Consequently, T_i is the timeout in iterative lookup set by $T_i \geq 2lt$. As a result, the expected latency of recursive lookup $E[RL]$ is defined in the following by these parameters.

$$E[RL] = (l+1)t + \frac{1-p^l}{p^l} T_r \qquad (2)$$

The expected latency of iterative lookup $E[IL]$ is also defined in the following.

$$E[IL] = 2lt + \frac{1-p}{p} lT_i \qquad (3)$$

In both recursive and iterative lookups, when l and t are fixed, p has a profound effect on performance. Figure 2 shows that an example of all expected latencies under different p when l and t are fixed values.
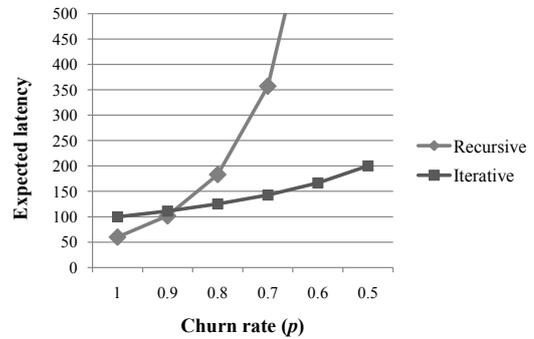
Figure 2: Expected latencies of recursive and iterative lookups under different p.

Moreover, T_r is much higher than T_i with this timeout setting. Thus, by using formula (2), the expected latency of recursive lookup increases especially when p is low. When p is low, on the other hand, iterative lookup does not have such high latency. However, when p is high, e.g. $p = 1$, this is higher than that of recursive lookup. Therefore, to increase out the performance of recursive and iterative lookups, we need determine what the system churn rate is.

2.3 Hierarchical DHT

Hierarchical DHT is a structure that divides a logical network configuration created by the DHT algorithm [8][9]. Figure 3 shows an example of a hierarchical DHT with two tiers in the Chord algorithm. Divided networks are called

top- and lower-level clusters. A top-level cluster is built by particular nodes called super nodes. Super nodes generally adopt strong nodes in the network, e.g., those with a great deal of high storage and high processing capacities that have been alive in the network for a long time, or those with wide bandwidth. Other normal nodes and specific a super node belong to lower-level cluster. The super node provides normal nodes with routes to other clusters.
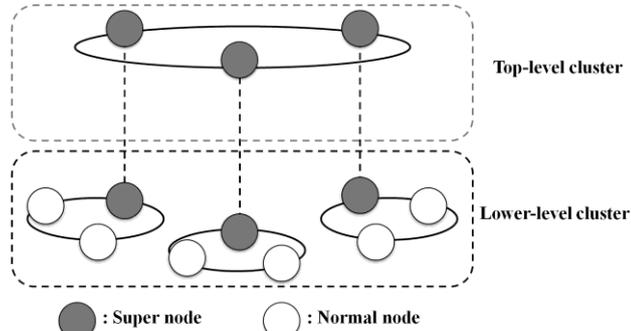
Figure 3: Example of two-tier hierarchical DHT.

Hierarchical DHT can speculate clusters where the destination of lookup belongs by comparing high m bits between the key and node ID. This m means the number of clusters in the hierarchical DHT by 2^m. When the high m bits of the key and a node ID are the same, the node forwards in the cluster. Otherwise, the node asks super node of the cluster to forward, and the super node finds the destination cluster and super node address by using the key.

Hierarchical DHT has various features, i.e., to assemble normal nodes as to their purpose and confine the effect of churn locally for neighbor nodes. An advanced study of hierarchical DHT found it to take into account the reliability of nodes [10]. This determines low-level clusters where normal nodes belong by using the interval from when they join to when they leave. The interval time is assumed by using a function, and this means that it is equivalent to R as life-time of a node. The function in this study assembled nodes that had similar R in each cluster. In addition, a super node was selected as a node that had the highest R in the cluster. Nodes are clusters obtained by R in this way in hierarchical DHT that takes reliability into consideration. Therefore, $E[R]$ becomes high due to clustering nodes that have higher R, and this also decreases by using clustering nodes that have lower R. Here, we assume that the interval for the stabilization process is fixed at all nodes and nodes obtain $E[S]$, which is almost a fixed value. p is defined as $E[R]$ in formula (1), and so this differs specifically for each cluster. Therefore, the p of each cluster can be speculated, and we can consider the optimal performance of a system that is appropriate to p.

3 GOAL AND PROBLEM

When p is low in recursive and iterative lookups, recursive lookup has an advantage, but iterative lookup has an advantage when p is high. We culled the lookups by using the churn rate. This ensured that the expected latency of lookups was the best under any churn rate. Our goal was to demonstrate this. To speculate churn rate p, we noted

hierarchical DHT took reliability into account. Hierarchical DHT determines clusters in which p is high or low as a result of clustering by the R of nodes. We focused on a structure where p was different for each cluster, and considered applying each lookup to that. However, each message format in recursive and iterative lookups uses differences for that. For this reason, a lookup cannot contact another lookup.

Here, we propose a strategy that changes over from one lookup to another by transforming the format of messages. We will explain how this strategy optimizes performance more than when only recursive or iterative lookup is used.

4 PROPOSED METHOD

4.1 System model

We propose that each cluster separates recursive and iterative lookups on hierarchical DHT to take reliability into consideration. We used the Chord algorithm because it had various features, e.g., it had a simple structure and was scalable. We also made note of the stabilization process for the reason of formula (3). Although super nodes were adopted in the clusters, we assumed that super nodes would be adopted in the system. This meant that the R of super nodes had no relationship to the R in the clusters. Here, the R of super nodes is R_s, and that of other normal nodes is R_n. Clusters in assembled nodes that have low R_n, called lower clusters, use iterative lookup in the clusters because they have low p. However, clusters in assembled nodes that have high R_n, called higher clusters, use recursive lookup. For example, a top-level cluster built by a super node has R_s. R_s is relatively high approximately R in the system. Therefore, a top-level cluster uses recursive lookup. There are recursive and iterative lookups in the system for this reason. Here, it transforms from recursive into iterative and vice versa about the message format. This process is executed at super nodes. This provides the communication between higher and lower clusters.

All nodes have a routing table built by the Chord algorithm to structure hierarchical DHT. For example, that of the normal node includes normal nodes that belong to the same cluster and super nodes of the cluster. Also, super nodes have routing tables that included normal nodes belonging to the cluster and the super nodes of the top-level cluster.

4.2 Transformed process

There are request and reply messages in recursive and iterative lookups. Each message format is different due to the lookup strategy. For example, a reply message including next hop candidates is used in iterative lookup as a routing table. However, no reply messages are used in recursive lookup. Tables 1 and 2 indicate that both request and reply messages have to include information at least in recursive and iterative lookups.

Table 1: Information in request message.

	Identifier	Key ID	Address of originator	TTL
Recursive	○	○	○	○
Iterative		○		

Table 2: Information in reply message.

	Identifier	Next hop Candidates
Recursive	○	
Iterative		○

Recursive lookup can forward in parallel because it trusts other nodes with forwarding request messages. Messages have to include the address of the originator, the message identifier to determine what value is received for which request message, and the Time To Live (TTL) which is set infinitely to forward request messages. The Identifier is set like the time made the request message. In iterative lookup, on the other hand, request messages do not have to include the address of the originator, identifier, or TTL because the originator controls the lookup. It only includes the key ID. However, reply messages must have some next hop candidates. Forwarding cannot continue because request and reply messages in both lookups are missing some necessary information.

By considering these differences, we implemented a transformed message format and lookup strategy. This transformed process particularly executes the transform from recursive to iterative and vice versa. It needs to be executed at all nodes on a flat DHT that does not have a hierarchy. However, the extent of the lookup strategy on hierarchical DHT is localized by clustering. For this reason, the transformed process is only executed at super nodes, which are contact points between clusters. The super nodes are confined to belong to lower clusters. They provide normal nodes with forwarding to top-level cluster and other clusters. Also, they provide other super nodes with forwarding to lower clusters. The flow for this operation of super nodes is outlined Figure 4.
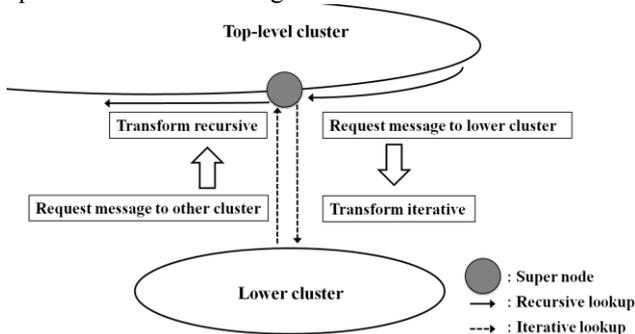
Figure 4: Transformed process at super node of Lower cluster.

When a super node receives a request message for iterative lookup from a normal node, if the destination is in another cluster, it creates a request message for recursive lookup from the subject matter of that message. However, the request message for iterative lookup does not include the identifier, the address of the originator, and TTL. For this reason, the super node creates a new identifier for the

request message by the time made the message, and sets the TTL from the route. Also, the address of the originator is specified by the super node. Normal nodes do not read messages for recursive lookup because they do not transform from recursive into iterative message format. Therefore, super nodes provide the originator with a forwarding destination node, and accept the reply message including the value with the transformed recursive into iterative message format.

However, when a super node belonging to a lower cluster receives a request message for recursive lookup, it can create a message for iterative lookup by only obtaining a key ID from the message. The value from the destination node similarly passes the super node, and it is sent the value of the transformed format.

4.3 Lookup strategy

We propose that higher clusters use recursive lookup, and lower clusters use iterative lookup. Here, a top-level cluster is recognized as a higher cluster and uses recursive lookup. As a result, the pattern for lookup executed in the above transformed process is categorized as two patterns, (A) from the lower to the top-level cluster, and (B) from the higher to the lower cluster.

First, Figure 5 shows an example of pattern (A).
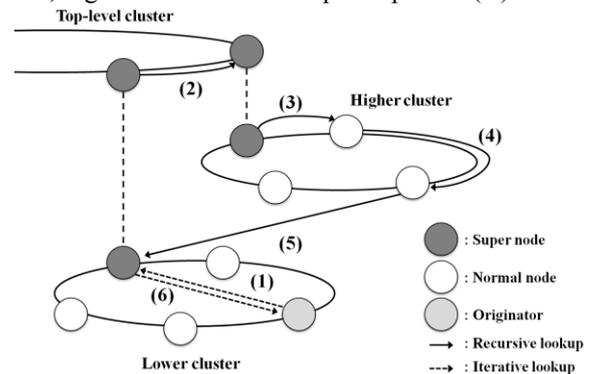
Figure 5: Lookup from lower to higher cluster.

The flow for lookup where request and reply messages are forwarded is indicated by the number in Figure 5. In addition, request messages for iterative lookup are transformed into those for recursive lookup. First, the originator requests a super node to forward to another cluster with iterative lookup (Figure 5 (1)). The super node transforms the message at the start, and starts recursive lookup. The lookup forwards to super and destination nodes (Figure 5 (2)-(4)). Although the destination does not directly send the value to the originator, it sends the super node belonging to the originator (Figure 5 (5)). The super node transforms the received message, and sends data to the originator (Figure 5 (6)). In this case, originator node waits for the message as Figure (6). However, the time may exceed the timeout of Iterative lookup. Here, we assume that super nodes do not leave the network, so some nodes certainly can communicate to super nodes. For the assumption, originator node waits to receive reply message from super node, because the node makes a reply certainly.

We consider that this pattern shorter the latency of the entire lookup more than that with only iterative lookup because it uses recursive lookup at the part with low churn.

Second, Figure 6 shows an example of pattern (B).
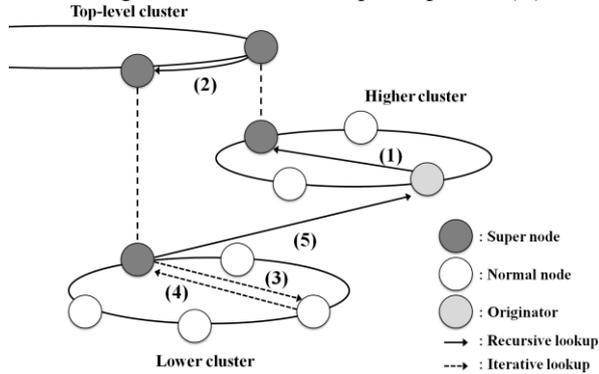
Figure 6: Lookup from higher to lower cluster.

A super node in this pattern executes the transformed process that creates a request message for iterative lookup from the request message for recursive lookup. Therefore, when the originator sends a request message for recursive lookup, lookup is executed at the super node of the destination cluster (Figure 6 (1), (2)). The super node executes the transformed process, and forwards destination by using iterative lookup (Figure 6 (3), (4)). The value is presented by using the communication shown in Figure 6 (4). The super node sends a reply message including the value for recursive lookup to the originator (Figure 6 (5)). Incidentally, the originator has to wait $2T_r$ because the lookup uses iterative lookup in the middle of lookup. By using iterative lookup at lower clusters where the churn rate is high, this pattern can shorten the latency of the entire lookup more than that with only recursive lookup.

5 EXPERIMENTS

5.1 Presupposition

We implemented the lookup in the Overlay Weaver [11] to evaluate our lookup strategy and compared its performance with that of only recursive or iterative lookup.

First, the setting for running the simulation and the version of the Overlay Weaver were:
- OS: Windows 7 Professional 64 bits
- CPU: Intel Core i5 3.2 GHz
- Memory: 4.0 GB
- Overlay Weaver: Ver. 0.10

Table 3 summarizes the parameters we set in the simulation.

Table 3: Parameters in simulation.

No. of nodes (N)	1000
No. of clusters (C)	4
Latency of one hop (t)	6 msec
Recursive timeout (T_r)	84 msec
Iterative timeout (T_i)	15 msec

C also means the number of super nodes, and C among N works as super nodes. Then, the lower-level cluster is built

by other nodes as normal nodes. Normal nodes have no relation to the distribution of R, and there is not much difference between the numbers of nodes in each cluster. T_r is based on the definition expressed in Subsection 2.2. We assumed that path length l was defined as $O(\log N')$ when N' was N/C as the number of one of the lower-level clusters. Also, we considered that it had the lookup of top-level clusters and a potential of over $O(\log N')$, and we added various values to l. T_r is defined by multiplying t by l. Similarly, T_i is multiplying t by 2 and adding a slight allowance because a node has to wait for a response in iterative lookup. Path length l is generally determined to be the key ID, which is a parameter that is not included in Table 3. This key ID is used the same as key ID to equalize the effect of l in all simulations as much as possible. By equalizing the effect, we ran the simulation for the key ID 100 times, and measured the average. In addition, we assumed that a higher and lower cluster were the same cluster in every simulation. We also assumed that churn rate p of higher clusters using recursive lookup was one at all times, and p in lower clusters using iterative lookup could be set freely. According to formula (1), p means the churn rate and needs S which is nearly a fixed value. For this reason, nodes repair fewer routing tables by churning during lookup. Additionally, the stabilization process was set to a large interval of 125 msec. This means $E[S]$ had a fixed value because nodes repaired fewer routing tables due to the stabilization process.

In addition, the following shows the routing tables of nodes.
- Predecessor node
- Successor List (not more than eight successor nodes)
- Finger table
- Normal nodes have super nodes in the cluster
- Super nodes have other super nodes in top-level cluster

When a normal node forwards a request message to another cluster, the node can forward the message to a super node in the same cluster in one hop. Additionally, a super node knows all of other super nodes in the lookup for the top-level cluster, and can forward the message to super node of the destination cluster in one hop.

We considered lookup where a normal node forwards request messages to the node of another cluster. Additionally, there are three lookup patterns for a cluster, and each lookup is executed in different nodes.

We measured latency from higher to lower clusters and otherwise with each lookup strategy using the above parameters.

5.2 Results

We measured average latency with simulation. Here, we assumed that the latency was the time until the destination node received a request message. In addition, the time also included the internal processing time of each node. Therefore, it measured $E[RL]$ and $E[IL]$ as follows in this simulation.

$$E[RL] = lt + \frac{1-p^l}{p^l}T_r \qquad (4)$$

$$E[IL] = 2(l-1)t + \frac{1-p}{p}lT_i \qquad (5)$$

First, we will consider pattern (B) in Subsection 4.3, which is a lookup whose destination cluster is higher. It assumes that the p of the higher cluster and that of the super node that belongs to a lower cluster is set to one at all times. Also, the originator does not leave the network. Additionally, we assumed that there was one lower cluster and three higher clusters. Therefore, we measured the average latency of nine lookup patterns that forward request messages to higher clusters. The results obtained from simulation are presented in Figure 7.
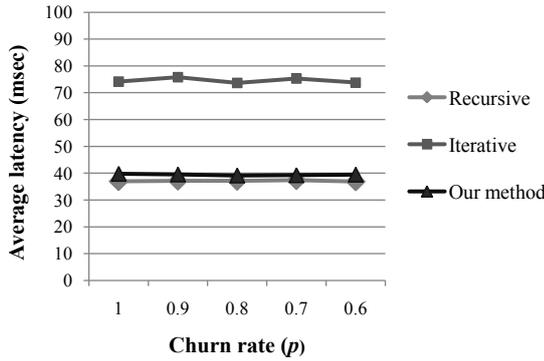
Figure 7: Average latency to three higher clusters by each lookup strategy.

This lookup pattern has little relevance to churn rate. First address is the super node belonging at the cluster because this lookup pattern necessarily forwards a request message to other cluster. The super node forwards the request message to super node belonging at destination cluster. Each node is assumed that churn does not occur. In addition, churn also does not occur after that because destination cluster is higher. As a result, the average latency hardly changes at all under any p. In Figure 7, when all nodes are steady state, the latency of our method is twice as short as that of iterative lookup. Although it is compared to recursive lookup, it has nearly latency of that.

Second, we will consider pattern (A) which is the lookup from higher to lower cluster. There are three lookup patterns from other three higher clusters. We set lower cluster to p which is single value from 1 to 0.6. Also, we ran a simulation for each p 100 times and measured the average latency in each lookup strategy. The result shows in Figure 8.

If churn increases in Figure 8, the average latency also increases. Recursive and iterative lookups are much the same as Figure 1. However, our method is same well as recursive lookup when p is one. If p decreases, increment of the average latency is similar to that of iterative lookup. Also, the result of our method is not identical with that of iterative lookup. Margin of average latency on each lookup is invariant from $p = 1$ to $p = 0.6$. Recursive lookup has the best average latency at only $p = 1$. However, from $p = 0.9$, recursive lookup has the worst average latency.
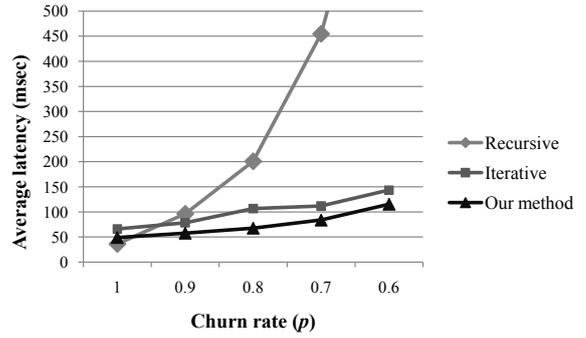
Figure 8: Average latency to one lower cluster by each lookup strategy.

Here, we will think expected latency of this structure. This means the latency when any node forwards. Also, this has relevance to the structure. For example, there are one lower cluster and three higher clusters in above simulation. If higher cluster is more than lower cluster, it is generally expected better latency. Because it is high probability that destination cluster is higher. On the other hand, if lower cluster is more than higher cluster, expected latency becomes low because it is high probability that destination cluster is lower.

For this reason, by these results, we measured the average latency of the structure. This was measured by multiplying each of average latency which destination cluster is both higher and lower by the number of higher or lower clusters. In this case, it multiplies result of Figure 7 by three as the number of higher clusters and that of Figure 8 by one as the number of lower clusters. Then, it measured the average of these results. We assumed that it is expected latency on the structure. Figure 9 shows the result of the case that there are one lower cluster and three higher clusters.
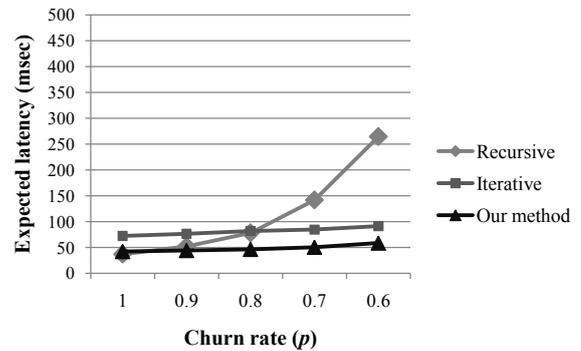
Figure 9: Expected latency on one lower cluster and three higher clusters.

This hierarchical DHT is made mostly of higher cluster, and so the expected latency is better than average latency to lower cluster. Iterative lookup and our method have flat latency well. Also, recursive lookup becomes better average latency than average latency of only lookup to lower cluster.

Here, we think about relationship between the average latency and the number of each cluster. In above case, we show the average latency that structure is one lower cluster and three higher clusters. We think that the average latency is influenced by the number of lower and higher clusters.

Therefore, we considered simulations which have different the number of these clusters within C.

First, we ran simulation that structure has two lower clusters and two higher clusters. Each lower cluster is set same p. In this case, we obtained six lookup patterns that destination cluster is lower. Also, there are six lookup patterns that destination cluster is higher. As it is for Figure 7 and Figure 8, we measured the average latency in each lookup pattern. Figure 10 and Figure 11 show each of average latency, to lower and higher cluster.
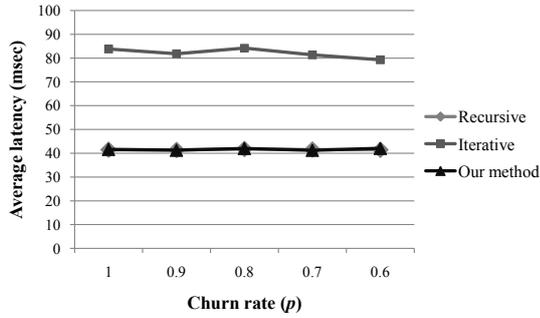
Figure 10: Average latency to two higher clusters by each lookup strategy.
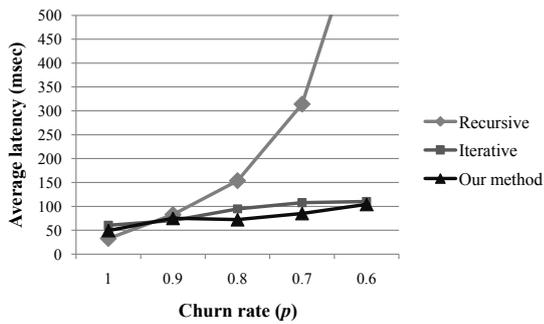
Figure 11: Average latency to two lower clusters by each lookup strategy.

These streams are not much more than Figure 7 and Figure 8. The result of Figure 10 is a little higher than that of Figure 7. Also, that of Figure 11 becomes low a little. However, these results are evaluated relatively, and they mostly equal. We will discuss minor margin about their data on Section 6. Similarly, by these results, we measure expected latency of this structure. The result is shown Figure 11.
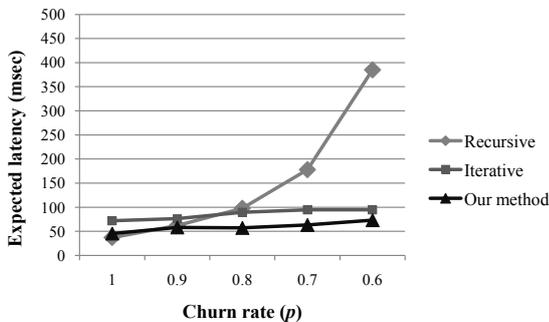
Figure 11: Expected latency on two lower clusters and two higher clusters.

This result is totally a little higher than result of Figure 9. When p is 0.8, the result of Figure 9 is that recursive lookup is lower than iterative lookup. However, Figure 11 shows that recursive lookup is higher than iterative lookup under the churn rate.

Second, we ran simulation that structure has three lower clusters and one higher cluster. In this case, lookup patterns that destination cluster is higher are three patterns. There are nine lookup patterns that destination cluster is lower. We measured the average latency each lookup pattern similarly. The average latency of the pattern that destination cluster is higher is shown as Figure 12. Also, we show the average latency to lower clusters in Figure 13.
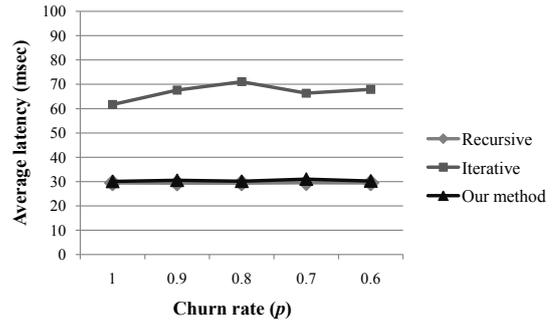
Figure 12: Average latency to one higher cluster by each lookup strategy.
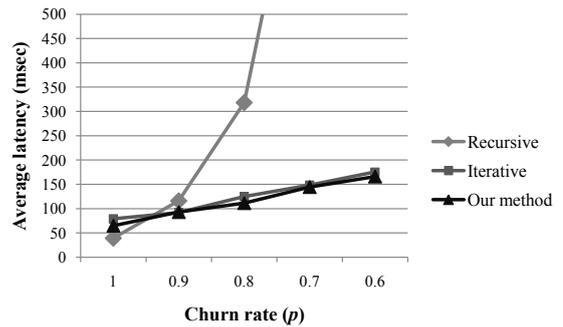
Figure 13: Average latency to three lower clusters by each lookup strategy.

These results have mostly same stream. However, max value of average latency to lower clusters is higher than other results to lower cluster. On the other hand, max value of average latency to higher cluster is better than other results. Similarly, by these results, we measure expected latency of this structure, and the result is shown Figure 14.
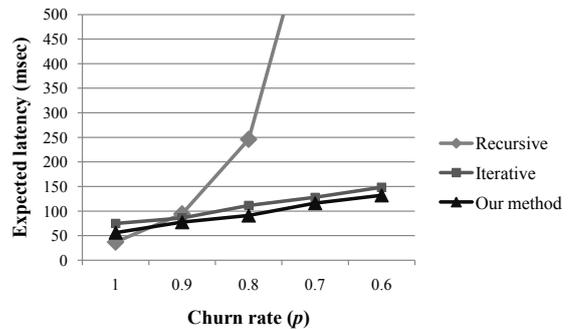
Figure 14: Expected latency on three lower clusters and one higher cluster.

This result is much higher than other results of expected latency. However, stream of the result is not much more than other results. In addition, when p is 0.9, the expected latency of recursive lookup is a little higher than other lookup strategy.

By these results, when nodes forward request messages to higher cluster, our method provides the performance of recursive lookup. Also, our method provided the performance of iterative lookup when nodes forward request messages to lower cluster. This is possible under any churn rate at lower cluster and proportion of higher to lower cluster. As a result, our method is effective when compared with only recursive or iterative lookup under any state and structure.

6   DISCUSSION

We will discuss about above results. First, we note effect of our method. In the case which destination cluster is higher, it has the similar performance of recursive lookup under any churn rate and structures. Also, when destination cluster is lower, it has the similar performance of iterative lookup under any situations. As a result, the expected latency of our method is relatively better than only other lookup strategy as integrated evaluation.

Second, we note the average latency of lookup to higher cluster on each structure. For the average latency of each lookup strategy, although the rate is almost same, the max value each of average latency is much different. This is considered that each lookup pattern have different path length. The path length is that it is five at minimum and eleven at maximum. If path length is long, latency becomes high. Therefore, average latency included the pattern had long path length becomes high. For this reason, the result of Figure 10 includes patterns had long path length, and that of Figure 12 does not include the patterns. However, the patterns also have reference to another lookup. For example, the result of Figure 13 becomes much high because it includes the patterns had long path length. However, we think that integrate effect of path length by measuring expected latency. If we consider effect of churn rate definitely, path length may have to be fixed.

For the results of expected latency, when higher cluster is defined $p = 0$, if super nodes know churn rate of each cluster and a number of clusters, we think that evaluate effective lookup strategy under the churn rate. For example, if p of a cluster becomes 0.8, the cluster uses iterative lookup when there are already two cluster using recursive lookup and one cluster using iterative lookup. This can know by Figure 11.

However, the case that p becomes 0 is less common in P2P. For this reason, we have to define higher and lower cluster. Therefore, we have to research about rigorous p and structure of clusters, a number of nodes and clusters.

7   CONCLUSION

We noted the effect of churn for recursive and iterative lookups in this study, and there were differences in the churn rate for each cluster on hierarchical DHT when the reliability of nodes was considered. We proposed a lookup method that will leverage both lookup advantages by culling the lookup strategy for each cluster. Additionally, we demonstrated that the new approach is significant in comparison to only recursive or iterative lookups. As a result, our method had the best expected latency under any churn rate. In future work, we need to consider an approach that dynamically applies our method to a DHT system. Additionally, we intend to propose an adaptive method that is able to adjust to variations in clusters by specifically defining the reliability of nodes and measuring the churn system. Also, we intend to consider various other parameters for the lookup strategy and how to provide optimal lookup.

REFERENCES

[1]   I.Stoica, R.Morris, D.Karger, M.Frans Kaashoek, and H.Balakrishnan, "Chord: A Scalable Peer-to-peer Lookup Service for Internet Applications", Proceedings of the SIGCOMM, 2001.

[2]   P.Maymounkov, D.Mazieres, "Kademlia: A Peer-to-peer Information System Based on the XOR Metric", IPTPS02, pp.53-65, 2002.

[3]   A,Rowstron and P.Druschel, "Pastry: Scalable, distributed object location and routing for large-scale peer-to-peer systems", IFIP/ACM Intermational Conference on Distributed Systems Platforms (Middleware), pp.329-350, 2001.

[4]   K.Shudo, D.Kato, Y.Kadobayashi, Y.Doi, "A Comparative Study of Iterative and Recursive Lookup Styles on Structured Overlays", The Special Interest Group Notes of IPSJ "system software and Operating System", Vol.86, pp.9-16, 2006.

[5]   S.Doi, S.Matsuura, K.Fujikawa, H.Sunahara, "Churn Tolerant Overlay Network Using Time Layered and Time Aggregation Methods", Transactions of Information Processing Society of Japan Vol. 51, No.4, pp.1142-1151, 2010.

[6]   S.Saroiu, "Measurement and analysis of Internet content delivery systems.", Doctoral Dissertation, 2004.

[7]   D.Wu, Y.Tian,K.W.Ng, "An analytical study on optimizing the lookup performance of distributed hash table systems under churn", Concurrency and Computation: Practice & Experience, Vol.19, pp.543-569, 2007.

[8]   L.Garces-Erice, E.W.Biersackm, P.A.Felber, K.W.Ross, and G.Urvoy-Keller, "Hierarchical Peer-to-peer Systems", ACM/IFIP International Coference on Parallel and Distributed Computing, 2003.

[9]   S.Zoels, Z.Despotovic, W.Kellerer, "On hierarchical DHT system - An analytical approach for optimal designs", Computer Communications, Vol.31, pp.576-590, 2008.

[10] F.Sato, "Configuration Method for Hierarchical DHT Systems Based on Join/Leave Ratio", Transactions of Information Processing Society of Japan Vol.51, No.2, pp.418-428, 2010.

[11] K.Shudo, "Overlay Weaver: An Overlay Construction Toolkit", HTML Available at, "http://overlayweaver.sourceforge.net/index-j.html".

Routing Metrics towards Reliable Communication in Mobile Ad-hoc Networks

Katsuya Fuji[†]and Takuya Yoshihiro[‡]

[†]Graduate School of Systems Engineering, Wakayama University, Japan
[‡]Faculty of Systems Engineering, Wakayama University, Japan
{s121048[†],tac[‡]}@sys.wakayama-u.ac.jp

Abstract - MANET (Mobile Ad-hoc Network) has been studied as future networks that are autonomously formed via wireless communication by each participating nodes. In MANET, links are frequently cut due to node mobility. In such case, it is desirable to change communication paths in advance to another reliable path before the communication path are disrupted. To this end, several dynamic metrics are proposed so far for link-state routing protocols such as OLSR. They raise link metrics when a link goes unstable to avoid using unstable links as communication paths between two nodes. Those dynamic metrics, however, involves routing loops that comes from inconsistency of transient network topology information (including link metrics) maintained at each nodes. This is so problematic since the routing loops cause severe congestion. In this paper, we propose a new routing metric that intends to prevent communication (service) disruption coming from node mobility, and simultaneously prevent routing loops coming from inconsistency of route computation. We evaluate the performance of our routing metrics through simulation and report the results.

Keywords: Ad-hoc Networks, Dynamic Metrics, Communication Reliability

1 Introduction

In recent years wireless communications have been populated and now wireless terminals such as smart phones are commonly used all over the world. One of the next step technologies, multi-hop wireless networks such as MANET have been focused on. In MANET, wireless links easily goes unstable and further node mobility often cut links between nodes. Therefore, how to realize reliable communication over such unstable links is one of the important issues in this area of study.

To make use of links as stable as possible at every time, several dynamic metrics are proposed so far [4][5][6][7][8][9]. From their results it is well understood that assigning low metrics to high-quality links makes communication reliability and throughput higher since high quality links are always selected as communication paths between nodes. They all try to maximize communication throughput. However, the factor to control throughput is too complicated so that we might not be able to control all of them with single approach such as dynamic metrics.

In this paper we assume that the congestion coming from interference, which is often said as one of the major reason

of link cut, is possible to be solved by other techniques, especially by MAC layer or Physical layer techniques. Namely, we focus on issues of IP layer. We concentrate on "link cut" due to node mobility, and "routing loops" due to transient inconsistency of network topology maintained at each node. By the assumption that there is no interference and physical obstacles, i.e., links are not cut by interference and physical obstacles, we can concentrate on those two factors (link cut and routing loops) that cause communication disruption. Note that even under this assumption it is still difficult to realize reliable communication. Our trial in this paper is to know whether dynamic metrics can ensure reliability of communication against those two IP-layer factors. Through investigating this point we try to get better understandings towards reliable communication over MANET.

More specifically, in this paper, we designed a new dynamic metric that is likely to work effectively against those two factors. In our metric design, we try to avoid link cuts due to node mobility by controlling link metrics sensitively according to the distance between nodes using RSSI (Received Signal Strength Indication). Further, to reduce routing loops, we apply methods called LMR[3] to suppress loops that limits metric range to change per unit time. We evaluate our dynamic metric with several mobility scenarios to show the potential of dynamic metrics on communication reliability.

This paper is organized as follows: In Sec.2, we shortly explain about related work on dynamic metrics on MANET. In Sec.3, we describe the design of our dynamic metric for communication reliability. In Sec.4, we evaluate our metrics through simulation, and finally in Sec.5 we conclude the work.

2 Related Work

2.1 Dynamic Metrics in MANET

In the literature, several dynamic metrics have been proposed. For wireless mesh network, in which nodes are stationary, one major metric is ETX (Expected Transmission Count)[4]. ETX is computed as success ratio of packet transmission between adjacent nodes based on hello packet receive ratio. ETT (Expected Transmission Time) [5] is developed as an extension of ETX, in which link speed is taken into account to estimate average transmission time as metric values. They compute metrics based on the number of hello messages received within a certain time period so that they have drawback that they are not sensitive enough to handle mobility scenario. Although there are several other dynamic metrics for wire-

less mesh networks, (e.g., MIC (Metrics of Interference and Channel Switching) [6] takes interference among nodes into account), they all have the same problem of sensitivity when we consider them to use in MANET with node mobility.

On the other hand, there are several routing metrics for MANET with mobility. If we consider a scenario where mobile nodes and non-mobile nodes are mixed, nodes with less mobility would be the nodes to expect reliable transmission. Based on this idea, Yawut et al. proposed a node metric that estimates amount of mobility of nodes [7]. In a scene where most of the nodes move around, however, relative speed between two nodes rather than speed of each node would be an essential metric. Tickoo et al. computes relative speed between two nodes using RSSI (Received Signal Strength Indication) [8]. Further, from the idea that the reliability of a path is determined by the least quality link, Triviño-Cabrera et al. proposed a path metric to compute a path with maximum least quality based on RSSI [9]. They all, however, do not consider transient routing loops, which is also an important factor to consider reliability of communication.

2.2 Dynamic Metric to Reduce Routing Loops

There are a few dynamic metrics in the literature that aware routing loops. The first loop-aware routing metric would be LLD [2], which constantly reduce link metrics little by little as time passes to prevent generating routing loops. LLD is based on the idea that the links with long living duration would be considered stable, so that link metric is designed depending on link duration. LLD, however, has a limitation that it cannot handle fluctuation of wireless link quality since the metric monotonically decreases as time passes.

As another loop aware method, LMR (Loop-free Metric Range)[3] is proposed. LMR can be applied to other dynamic metric scheme to reduce routing loops by limiting the amount of metric change per unit time. LMR defines the variable $r(> 1.0)$ called *metric stretch* to limit the range of the next metric value by r in ratio. Specifically, let m_{old} be the current metric of a link and m_{new} be the next metric, then $m_{old} \cdot r^{-1} < m_{new} < m_{old} \cdot r$ must be held. If the new metric of the base dynamic metric scheme is out of the range, LMR changes the new metric with the nearest value in the range. LMR is illustrated in Fig. 1; the base metrics are changed within the range of LMR to slow the transition of metrics.

In the analytical results of LMR[3], it is reported that the value of r to guarantee loop-freeness under the assumption that no control packet drops can be calculated. Note that the value of r to guarantee loop-freeness depends on several values: the minimum link metric value m_{min}, the maximum one m_{max} and the network diameter w in hop count. Consequently, once m_{min}, m_{max} and w are given, the value of r to guarantee loop-freeness under no control packet loss condition.

Unfortunately, however, the value of r to guarantee loop-freeness is too small in general (e.g., the value is 1.002 when $m_{min} = 1, m_{max} = 5$ and $w = 10$[3]), so we have to use larger value. Note that when r gets smaller, we can expect more effects on reducing loops whereas the effects of base dynamic metric scheme is limited since the dynamism of met-
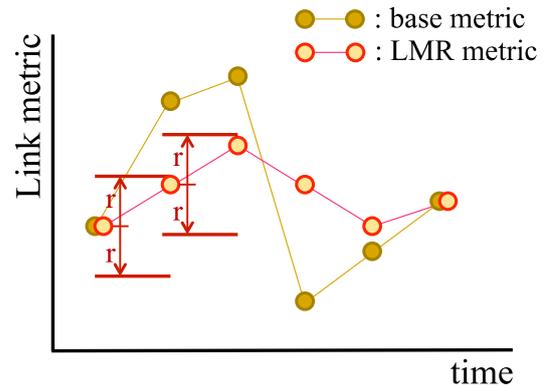
Figure 1: LMR: the method to reduce routing loops

ric change is weakened. There is the trade-off between them so that we have to pay attention to the balance to work with optimal performance.

3 Design of Routing Metrics for Reliable Communication

3.1 The Concept

As we mentioned before, we designed a dynamic metric that aware both link disruption due to node mobility and routing loops due to route inconsistency. In this paper, we suppose proactive link-state routing scheme such as OLSR[1] as a base routing scheme.

To prevent communication disruption coming from link cut, we have to raise the metric of a link high enough to have detour paths available before the link is cut due to node mobility. To this end, the routing metric must be sensitive to detect the symptom of link cut; it is not ideal to use packet loss statistics like ETX since that approach requires long-term (at least about 20 seconds) observation. In our metric design, we use RSSI (Received Signal Strength Indicator) values measured on hello messages transmitted by neighbor nodes as a sensitive routing metric. Specifically, we intend to estimate the distance between two nodes, and the link metric is computed according to the distance. Note that the distance estimation algorithm we use is quite primitive so that they might work surely only in simulation. Note that, however, we can use more practical distance estimation algorithms such as [10].

To prevent routing loops, we apply LMR into our RSSI based routing metric. Note that LMR limits the effect of dynamic metrics instead of reducing routing loops. To take balance of them, we have to select carefully the value of r.

3.2 Design of RSSI based Dynamic Metrics

The formula that indicates relation between distance and metric should be carefully designed. See Fig.3.2. Fig.3.2(a) is a case in which distance between node A and C is about to reach the communication range. In this case, $A \rightarrow B \rightarrow C$ is the desirable communication path from A to C since link (A, C) is about to be broken. This indicates that the condition $m(A, C) > m(A, B) + m(B, C)$ should be stand,

where $m(A, C)$ is the metric of the link (A, C). On the other side, Fig.3.2(b) shows the case where the shape of node location is the same but the distance between A and C is still shorter than the communication range. In this case, the desirable communication paths from A to C is $A \rightarrow C$ to reduce hop count that packets travel. This indicates the condition $m(A, C) > m(A, B) + m(B, C)$ in turn.

Now we see that the desired paths differ even if the shape of node location is the same, and also that selecting detour paths is more desirable as the distance become longer. From this consideration, we conclude that the metric function should be convex function, if we consider continuous metric function. In this work, we use a simple polynomial convex function as a metric function.

Now we in turn mention about mechanisms to generate metrics. The distance between nodes is estimated from RSSI observed on every hello messages of OLSR. Namely, when a hello message is received, the metric of the directed link to the neighbor is computed. The metrics are propagated via hello messages and topology advertisement messages (TC messages in OLSR) all over the network to be used in shortest path computation.

Specifically, the RSSI values of hello messages are used to estimate distance between nodes. As we assume no physical obstacles, RSSI decays in inverse relation to square distance. So, let L be the distance between two nodes, R be RSSI measurement and a be decay coefficient, then we have,

$$L = \sqrt{\frac{a}{R}}.$$

Note that the constant a should be determined properly.

Then we compute the routing metric from the estimated distance. The metric function we use is a simple polynomial function as mentioned before. Let L_{max} be the maximum distance to communicate, M_{min} be the minimum metric, M_{max} be the maximum metric, then the function is expressed as follows:

$$M = (M_{max} - M_{min}) \times (\frac{L}{L_{max}})^n + M_{min},$$

where n is a parameter to determine the curve of the function. Fig.3 shows the curve of the metric function for several value of n. You see that the metric value take values between M_{max} and M_{min} when the distance less than L_{max} is given.

3.3 Applying LMR

To reduce routing loops we apply LMR to our RSSI based metrics shown in Sec.3.2. Namely, by LMR, we limit the amount of metric change using metric stretch r. Note that the amount of metric change in our RSSI based metric depends on relative speed, moving direction and distance between nodes. If limitation of LMR is too strong compared to nodes' speed, links are cut before their metrics raise enough to change communication paths, resulting in communication (service) disruption. Conversely, if limitation of LMR is too weak, routing loops appear to lead severe congestion. It is important to take this balance.
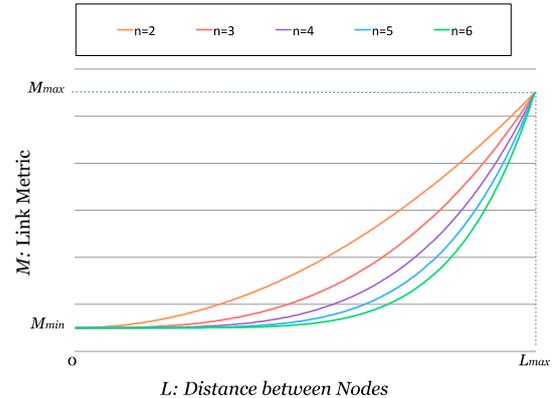
Figure 3: The Curve of the Proposed Metrics

We also note that that our metric will not work when node speed is too fast. To find the node speed that our metric can catch up with is also important in evaluation.

4 Traffic Simulation

4.1 Scenario

We conducted a simulation using network simulator Qualnet[12]. We implemented both our RSSI based routing metric and LMR by modifying OLSR module OLSRv2-NIIGATA, which is included in Qualnet version 5.0. We compared the performance of (i) the RSSI based metric and (ii) the RSSI based metric with LMR in the simulation.

As simulation scenario, we prepare 1000m×1000m field to place 30 nodes in random location. Nodes move to follow Random Way Point model [11]. We try two sort of node speed, i.e., 5km/h and 10km/h, and the pause time of nodes is 10 seconds. We generate 5 flows of 20kbps CBR (Constant Bit Rate) in 5 minutes, i.e., start at 1 minute and ends at 6 minutes from the beginning of the simulation. As the parameter that determine the curve of metric, we use $n = 4$. Note that OLSR have the mechanism called MPR to reduce control message load. To exclude the effect of MPR, we set TC_REDUNDANCY=2 so that all links are propagated into the network. As other OLSR parameters, we use default values. When 1 minute past from the beginning of the simulation, all nodes share the information of all links.

4.2 Results

In Fig.4, we show the packet reachability on destinations for various values of r. Here, "power" means the case of RSSI metrics without LMR. In both 5km/h and 10km/h, the performance gets worse as the value of r goes lower. Totally the performance of lower mobility (5km/h) is better than higher mobility (10km/h).

In Fig.5, we count the number of loop packets. We define loop packets as the packets reaching the same node more than twice. First we found the number of loop packets is much larger in high mobility than low mobility. Note that there are many loop packets especially when $r = 1.01$ in lower mobility. This would be the effect of link cuts; when the lim-

(a) Case of A and C having long distance   (b) Case of A and B having relatively small distance
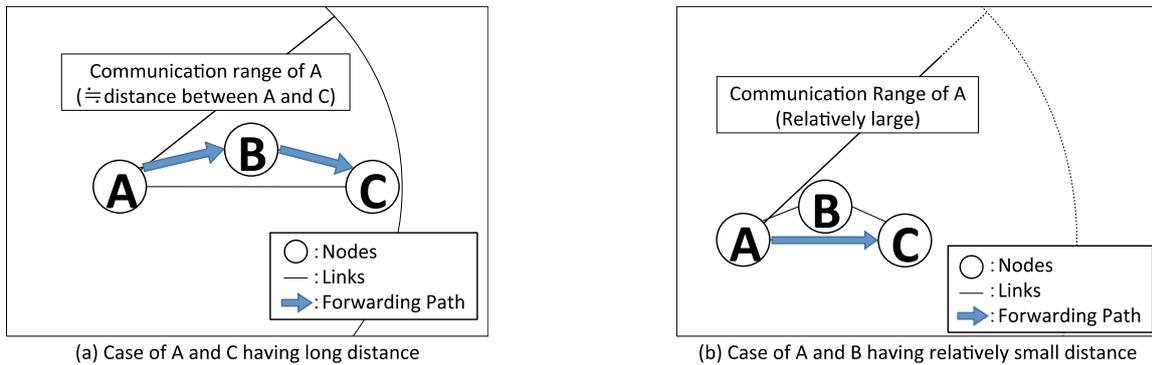
Figure 2: The Desirable Path with node distance
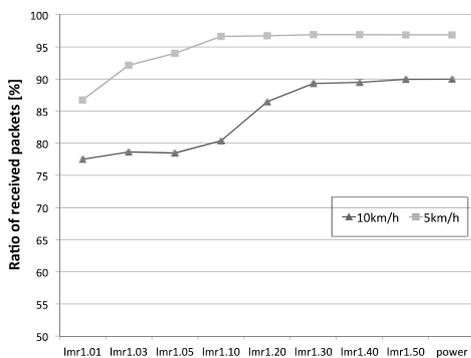
Figure 4: Packet reachability
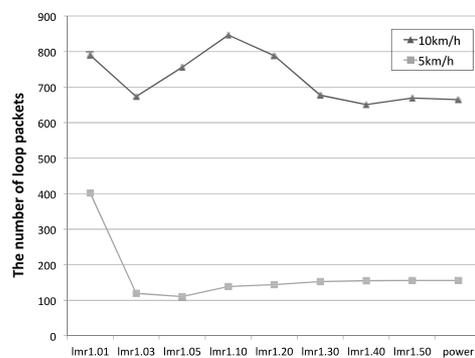
Figure 5: The number of loop packets

itation of LMR is too strong, links are cut before its metric gets raised to enough value, and consequently if its new next hop obtained from path recomputation is its previous hop in the old paths, the packet loops. Link cuts generate routing loops in this way and is seen in Fig.5. Also, in low mobility case, the number of loops gradually raises as r goes larger from 1.05. This implies that the effect of LMR appears when r gets lower. But when R is lower than 1.03 the effect of link cut exceeds to this loop reduction effect. As a result, we conclude that the balance point is seen at $r = 1.05$ in lower mobility scenario.

In contrast, such balance point is not seen in higher mobility scenario. This is because the effect of link cut is seen in higher r, while that of routing loops is seen in lower r, and in the middle r they are mixed. At the point around $r = 1.10$, the situation is the worst since the bad effect of both factors gather and their synergetic effect appears. In fact, when we checked the event log of the simulation, the chain of those two factors, i.e., link cut cause loops then the loops cause another link cut and so on, are seen frequently.

In Fig.6 and 7, we show the number of flow cut and the total flow-cut time in both 5km/h and 10km/h scenario. We regard "flow cut" if CBR packets are not received in a period of more than 2 seconds at the destination node of each flow. The flow-cut time is the total sum of all link-cut duration of all of 5 CBR flows. Regardless of node speed, the performance

goes worse when r goes smaller.

4.3 discussion

In Fig.4, the difference of packet reachability between $r = 1.01$ and $r = 1.03$ in 5km/h scenario is about 5%, which is about 400 packets. From Fig.5, the difference in loop packets is about 300 packets. It implies that 3/4 of the packet loss is caused by packet loops. On the other hand, in Fig.6, link cuts get raised when r is lower than 1.05, implying that the effect of link cut appears in packet loops only when the degree of the effect is strong.

As another findings, see Fig.5: we see the peak at $r = 1.10$ in 10km/h scenario. If we see the right side slope of the peak and packet reachability (Fig.4) at the same r, we find that the packet reachability rises as loop reduces. If we see the left side slope, however, packet reachability does not change although loop packet reduces. This implies that packets are merely dropped instead of looping around $r = 1.05$.

Finally we mention about the balancing point at $r = 1.05$ in 5km/h scenario. As the balance point is seen in 5km/h scenario, we can separate the effect of link cut and routing loops; if $r < 1.03$ the effect of link cut strongly appears while if $r > 1.05$ the routing loops and loop reduction effect of LMR appear. This makes us expect that in slower mobility scenario the effect might be seen clearly. Note that if message interval of OLSR is set smaller, it is the same as we lower the
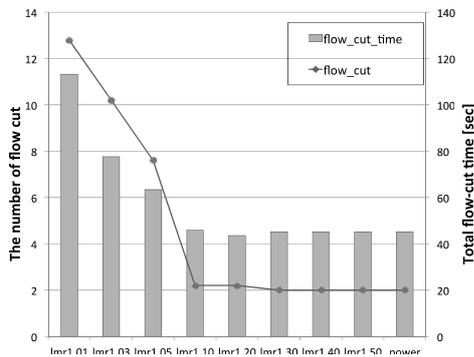
Figure 6: Flow cuts in 5km/h scenario
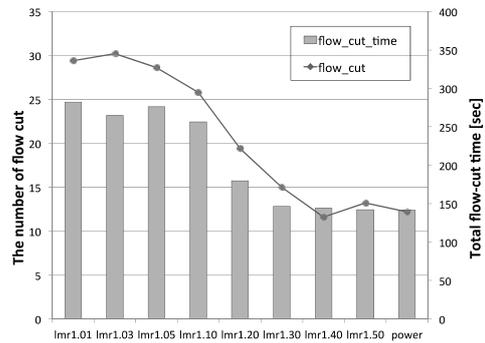
Figure 7: Flow cuts in 10km/h scenario

mobility. As a first step we expect to support 5km/h mobility (which is walking speed), trying smaller message interval will be expected in the next step.

5 Conclusions

In this paper we designed the new dynamic metric to solve the problem of communication (service) disruption caused by link cut due to mobility and routing loops. Our metric is designed using RSSI and LMR to solve both factors.

Through simulation using Qualnet, we found our dynamic metric works effectively in 5km/h scenario. We also found the balance point of r in performance. In contrast, in 10km/h scenario the effect of link cuts and routing loops prevent our dynamic metric from working well.

As future work, we would like to try the scenario of lower mobility or smaller message interval scenario for further understandings.

Acknowledgment

REFERENCES

[1] T. Clausen and P. Jacquet, "Optimized Link State Routing Protocol (OLSR)," RFC3626, 2003.

[2] Takuya Yoshihiro, "LLD: Loop-free Link Stability Metrics for Proactive Link-State Routing in Wireless Ad Hoc Networks," Mobile Ad-Hoc Networks: Protocol Design, Xin Wang (Ed.), ISBN: 978-953-307-402-3, InTech, 2011, Available from: http://www.intechopen.com/articles/show/title/lld-loop-free-link-stability-metrics-for-proactive-link-state-routing-in-wireless-ad-hoc-networks.

[3] T. Yoshihiro, "Reducing Routing Loops under Dynamic Metrics in Wireless Mesh Networks," IEEE Global Communication Conference, Exhibition, and Industry Forum (Globecom2010), pp.1-6, 2010.

[4] DeCouto, D.,Aguayo, D., Bicket, J. Morris,R. A High-Throughput Path Metric for Multi-Hop Wireless Routing, Proceeding of ACM Annual International Conference on Mobile Computing and Networks (MOBICOM2003), pp.134—146, 2003.

[5] Draves, R., Padhye, J. and Zill, B. Routing in Multi-Radio, Multi-Hop Wireless Mesh Networks, Proceedings of ACM Annual International Conference on Mobile Computing and Networks (MOBICOM2004), pp.114—128, 2004.

[6] Yang, Y., Wang. J. and Kravets, R., Interference-aware Load Balancing for Multihop Wireless Networks, In Technical Report UIUCDCS-R-2005-2526, Department of Computer Science, University of Illinois, 2005.

[7] Yawut, C., Paillassa, B. and Dhaou, R. (2007). On Metrics for Mobility Oriented Self Adaptive Protocols, Proceedings of in Wireless and Mobile Communications 2007 (ICWMC2007), 2007.

[8] Tickoo, O., Raghunath, S. Kalyanaraman, S. Route Fragility: A Novel Metric for Route Selection in Mobile Ad Hoc Networks, Proceedings of IEEE ICON03, pp.537—542, 2003.

[9] Triviño-Cabrera, A., Nieves-Perez, I., Casilari, E. and Gonzalez-Canete, F.J. Ad Hoc Routing Based on the Stability of Routes, Proceedings of the 4th ACM International Workshop on Mobility Management and Wireless Access, pp.100—103, 2006.

[10] Y. Shang, W. Ruml, Y. Zhang, and M.P.J. Fromherz. Localization from mere connectivity. In Proc. ACM MobiHoc, pages 201—212, Annapolis, Maryland, USA, Jun. 2003.

[11] Broch, J; Maltz DA, Johnson DB, Hu Y-C, and Jetcheva J (1998). "A performance comparison of multi-hop wireless ad hoc network routing protocols," proceedings of the Fourth Annual ACM/IEEE International Conference on Mobile Computing and Networking(Mobicom98), ACM, October 1998.

[12] Qualnet, http://www.scalable-networks.com/products/qualnet/ .

A Study on Stream Prediction Based on Timing Correlation among Multiple Data Stream

Kenji Arai[†], Yoh Shiraishi[‡], and Osamu Takahashi[‡]

[†] Graduate School of Systems Information Science, Future University Hakodate, Japan
[‡] School of Systems Information Science, Future University Hakodate, Japan
{g2110002[†], siraisi[‡], osamu[‡]}@fun.ac.jp

Abstract Research on data streams has attracted a great deal of attention in many fields in recent years, such as sensor-network technologies and stock quote data. Therefore, stream prediction technologies have attracted the attention of stream mining technologies. When we want to obtain the predicted value of a certain single data stream, most methods use past data on the data stream. However, we think that correlations, such as synchronization, can be used for method of predicting streams, and their accuracy might be better than methods that only use past data on single data streams. In addition, we need to take into consideration that correlations are not based on synchronization, which we call "similarity correlation". We suggest a method of detecting "timing correlation" from multiple data streams in this paper, and a method of predicting streams based on these correlations. We also discuss our demonstration of the efficacy of these methods.

Keywords: Data Stream, Clustering, Classification, Correlation, Prediction

1 INTRODUCTION

Research on data streams has attracted a great deal of attention in numerous fields in recent years, such as sensor-network technologies and stock quote data. Data streams are expressed as time-series data of unlimited length and they increase in real time. Stream mining technologies have been studied intensively [1] to find significant patterns from data streams. For example, some researchers have studied about trends in data streams [2]. The trends in data streams are diverse. They express a lot of features, such as periodicity. They are used to predict future data of data streams according to trends of past to present.

When a certain data stream is predicted, most methods only use past data on the predicted object. However, if two or more data streams are measured from sensors installed indoors, their correlation, such as synchronization of values, might appear from these data streams. We think, their correlations can be used by methods of predicting streams, and their accuracy might be better than approaches that only use past data on single data stream. We assumed the situation where two data streams were measured simultaneously in Figure 1 as an example of this theory. The wave-forms of temperature data change according to humidity data near the current time. The accuracy of prediction may decrease when strange values like these are measured, if we only use a method of prediction using past

data on single data streams. However, if the correlation that temperature data increase after a rapid increment in Humidity has been found, accuracy of prediction will be improved by taking into consideration this correlation.

We assumed we needed to consider two correlations between data streams.
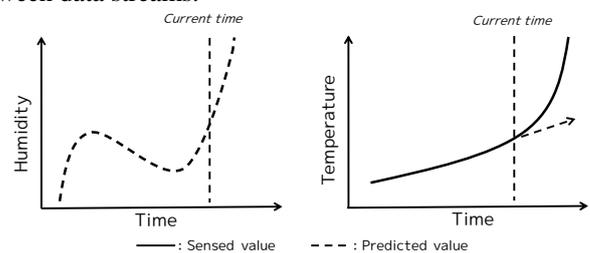
Figure 1: Sample of Stream Prediction Based on Correlation

Ⅰ. Similarity Correlation

Streams A and B in Figure 2 are measured simultaneously. The two data streams are similar because of their measured values and their trends are very close. There is a correlation based on similarity between these two data streams. We call this correlation "similarity correlation" in this paper.
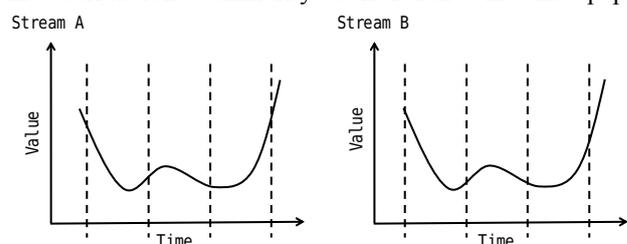
Figure 2: Example of Similarity Correlation

Ⅱ. Timing Correlation

Streams C and D in Figure 3 are measured simultaneously.
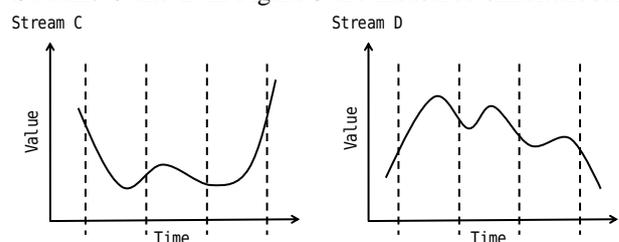
Figure 3: Example of Timing Correlation

The measured value and trend of stream C differ from those of stream D. However, we assumed there was a rule based on the timing of changing trends. For example, when

stream C increase, stream D decrease. We found this rule to be one of correlation and call it "timing correlation" in this paper.

Our task was to predict data streams based on similarity and timing correlation, and we solved this by using our new approach.

2 RELATED WORK

2.1 Related Work in Stream Mining

Many researchers in stream mining technologies have studied trend detection technologies, such as detection trends from partial sequences made from data streams [3][4]. Trends represent features, such as the periodicity of partial sequences and data streams. Kawashima et al. found a method of reducing the number of calculations by cutting off the dimensions of partial sequences with adaptive piecewise constant approximation [5]. They succeeded in fast matching of partial sequences to sample sequences in the database with this method. Toyota et al. introduced a method of detecting trends method based on the dynamic time warping (DTW) distance [6]. Papadimitriou et al. [7] predicted future data of data streams by using models like the auto regression model using the wavelet coefficients of data streams. The coefficients of wavelet transforms and model updates are executed each time a stream data is detected at minimum cost because previous results are used. However, this method does not take into consideration the correlation between multiple data streams.

In addition, Sakurai et al. introduced a correlation detection technology called "BRAID" [8]. Zhu created StatStream [9] to detect correlations by comparing the DFT coefficients of partial sequences. Although, these methods took into consideration similarity correlation, timing correlation was not considered. Moreover, they did not refer to stream prediction.

2.2 Relationship with Data Mining

Data mining technologies are similar to those for stream mining. They find a significant rule from time-series data stored at databases. Stream mining technologies are regarded as one field in data mining technologies. However, stream mining technologies must take into consideration the amount of memory and processing time, because data streams is time-series data that increase in certain interval in real time. Therefore, some methods in stream mining technologies are different from them of data mining technologies. In stream mining technologies, the processing time and amount of memory are more important than complete result. In addition, approximated solutions are generically used instead of exact solutions at real time processes. Therefore, incremental algorithms are valued because they decrease the number of calculation. These algorithms use previous results to calculate new results.

3 PROPOSED METHOD

3.1 Overview of Proposed Method

We explain the requirements for our approach to solve our tasks in this section.

Data streams are predicted from estimation of future trends by taking into past trends in data streams. Consequently, a method of detecting trends and correlations from past measured data is required. In addition, the method of detecting correlations need to be based on numerous features of data streams to detect timing correlations, and it cannot use degrees of similarity such as the DFT coefficient and DTW distance. Therefore, these methods need to be incremental and only use a certain amount of memory.

First, we will explain the environment for data streams in our approach. Two or more data streams are measured simultaneously. All data in the data streams are measured accuracy at certain intervals, no missing or delayed data.

Our method of detecting correlations and predicting data streams is clarified in what follows. Our method detects trends from data stream in real time. Next, it manages information on the appearance of trends. Correlation detection in data streams is provided by matching detected trends and the number of relations base on past trends. After that, our method predicts future data stream according to correlations between data streams. It:

1. Detects new trends.
2. Manages information on the appearance of trends.
3. Detects correlations between data streams, and
4. Predicts future data stream based on correlations.

3.2 Detection of Trends

A) System for Classifying Trends

Our method divides data streams into partial sequences in real time, and detects trends in partial sequences as current trends in the data streams by using a classification system. Therefore, it is necessary to prepare a classification system from past data. The method executes five steps to construct a classification system. It:

i. Divides all data streams into partial sequences by a certain interval.
ii. Extracts feature quantity patterns FQ-P and FQ-N from all partial sequences.
iii. Clustering partial sequences according to FQ-P.
iv. Clustering partial sequences associated with each cluster according to FQ-N, and
v. Constructs a classifier from the clusters.

a) Divide Streams into partial sequences

Our method divides each data stream into multiple partial sequences with a certain interval. Each interval is called a window. We have assumed partial sequences divided by the same window have the same number of data. All data streams are divided by the same number of partial sequences.

b) Extract Feature Quantity Patterns

The method extracts feature quantities that express trends in all partial sequences, and bundles a number of feature

quantities as a pattern. This pattern is called a feature quantity pattern in this paper.

We used the power spectrum obtained from the DFT coefficients of partial sequences and natural values of partial sequences as resources for feature quantities. We called a feature quantity pattern obtained from the former an "FQ-P". We also called the latter an "FQ-N". The method extracts follows.

- Maximum Value / Data index of maximum value
- Minimum Value / Data index of minimum value
- Variance
- Value of integral

FQ-P expresses the periodicity of partial sequences while FQ-N expresses other waviness features.

c) Clustering Partial Sequences

Our method classifies partial sequences into various groups (clusters) according to the extracted feature quantities (feature quantity pattern). A wide variety of feature quantity patterns can be generated from partial sequences. Therefore, it is difficult to detect the number of clusters in advance. We must have used clustering algorithm that could automatically detect the number of clusters. In addition, the clustering algorithm must take into consideration similarity of classified partial sequences in a cluster, such as those in Figure 4.
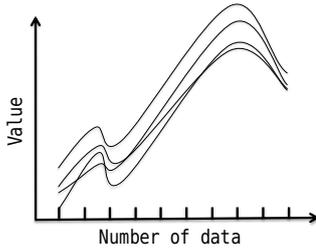
Figure 4: Example of Partial Sequences in Cluster

Therefore, we devised a clustering algorithm that uses a cluster division standard based on Eqs. (1) and (2). In these equations, variable k is the number of cluster, variable n is the number of partial sequences in cluster k, and variable m is the number of feature quantities in a feature quantity pattern. Variable x is the variable j th value in the i th partial sequences in cluster k. Variable V_{kj} is the variance of the j th values obtained from all partial sequences in cluster k. Variance E is the standard deviation of V_{kj}, and this is the cluster division standard. The variable E of a cluster expresses the variability of partial sequences.

$$V_{kj} = \sum_{i=1}^{n} x_{kij}^2 / n - \left(\sum_{i=1}^{n} x_{kij} / n \right)^2 \qquad (1)$$

$$E_k = \sqrt{\sum_{j=1}^{m} V_{kj}^2 / m - \left(\sum_{j=1}^{m} V_{kj} / m \right)^2} \qquad (2)$$

Further, the method uses two feature quantity patterns to classify strictly partial sequences into clusters according to similarity of multi-aspects. Firstly, it uses FQ-P for classification. Secondly, it classifies partial sequences in

each cluster into new clusters according to FQ-N. The clustering algorithm using cluster division standard E involves six steps.

i. Classify partial sequences into two clusters by two-means clustering according to FQ-P, and add these clusters to the cluster list.
ii. Classify each cluster into two clusters by two-means clustering. The divided clusters are the "parent cluster", and the new clusters are the "child clusters".
iii. Calculate each cluster's division standard.
iv. If the average of a child cluster's cluster division standard is lower than that of the parent cluster, add child clusters to the cluster list and remove the parent cluster.
v. Repeat processes 2 to 4 while the clusters are divided.
vi. Repeat processes 2 to 5 according to FQ-N

d) Construction of classification system

The method constructs a classification system, to classify partial sequences in real time. A classification system is constructed for all data streams. We used C4.5, which is an algorithm for constructing decision trees implemented in Weka [10] as J48. All feature quantity patterns in each cluster are used for the construction. A decision tree classifies a partial sequence into a cluster according to the thresholds of feature quantities, as shown in Figure 5, which is created from the feature quantity patterns of stream A.
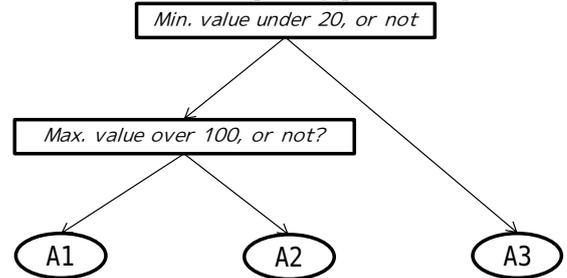
Figure 5: Example of Decision Tree for Feature Quantities

B) Detect Trends using Classification system

Our method detects trends from new partial sequences measured in real time. Each partial sequence obtains a cluster number. These number enable to express partial sequences as cluster number streams, as shown in Figure 6, where the partial sequences of data stream A are classified into {A1, A2, A3}.
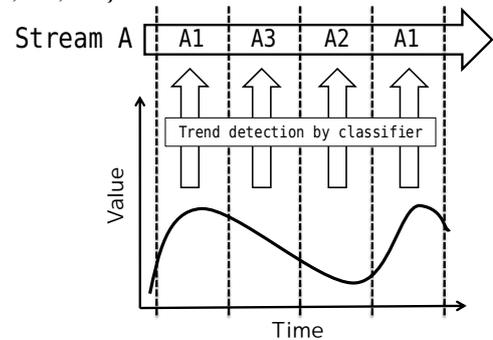
Figure 6: Example of Trend Detection in Stream A

Because the first and fourth partial sequences are similar, the system for stream A classifies both partial sequences

into the same cluster, such as A1 in Figure 6. Other partial sequences are classified into each cluster.

3.3 Managing Correlations

Our method detects the correlation between information on the appearance of trends. Our method manages two kinds of correlation rule between trends. The first is for the correlations between trends from different kinds of data streams. The second is for the correlations between trends in continuing windows from each data stream. We introduced a cluster correlation table to manage the former correlations and a cluster transition table to manage the latter correlations.

A) Cluster Correlation Table

The method executes trend detection at certain intervals, and a cluster simultaneously appears in all data streams. The cluster correlation table, in Figure 5, manages the relations between clusters from different data streams that appear. Stream A and B have been measured simultaneously in this figure. Partial sequences from streams A, and B are classified into {A1, A2, A3}, and {B1, B2, B3}. For example, when a partial sequence in stream A and partial sequences in stream B are simultaneously classified into A1 for the first and B3 for the second, corresponding cell (A1, B2) is updated. Each cell means the frequency that clusters from different data streams appear simultaneously.

		Stream A			Stream B		
		A1	A2	A3	B1	B2	B3
Stream A	A1				5	2	1
	A2				20	20	15
	A3				9	8	38
Stream B	B1	11	20	9			
	B2	29	20	8			
	B3	15	15	38			

Figure 7: Sample of Cluster Correlation Table

Transition rule		Count
From	To	
A1	A1	12
A1	A2	1
A1	A3	33
A2	A1	19
A2	A2	11
A2	A3	4
A3	A1	32
A3	A2	4
A3	A3	5

Figure 8: Example of Cluster Transition Table

B) Cluster Transition Table

We use a cluster transition table to estimate a cluster for a partial sequence in the next window of a stream. A cluster transition table manages the cluster transition rules for each stream. Figure 8 has an example of a cluster transition table for stream A. When the cluster of the targeted window is A1 and the cluster of the next window is A2, the transition rule is A1→A2. The number of transitions in the corresponding record is updated every time a transition occurs. By referring to this cluster transition table, we can examine a cluster that is easy to change from the targeted cluster. For example, since there are many transitions of A1 →A3 and A3 →A1 in Figure 8, these transitions are likely to occur.

3.4 Prediction Based on Correlation

A) Prediction Cluster based on Correlation

The method predicts a trend of the next partial sequence based on correlations between clusters. It involves the four steps in Figure 9.

i. First, it determines the latest window according to the current time. The current time exceeds the end time for the latest window. The current window, including the current time, does not yet have enough data to detect trends with a classifier made for the target stream.

ii. Next, it searches a cluster correlation table for a correlated cluster that is likely to occur simultaneously with the latest cluster from the targeted stream.

iii. It searches the cluster transition table of the stream including the correlated cluster for a cluster (estimated cluster) that will occur in the next window.

iv. Finally, it again searches the correlation table for the cluster of the targeted stream that is likely to occur simultaneously with the estimated cluster. The cluster that is searched is the result of this prediction process.
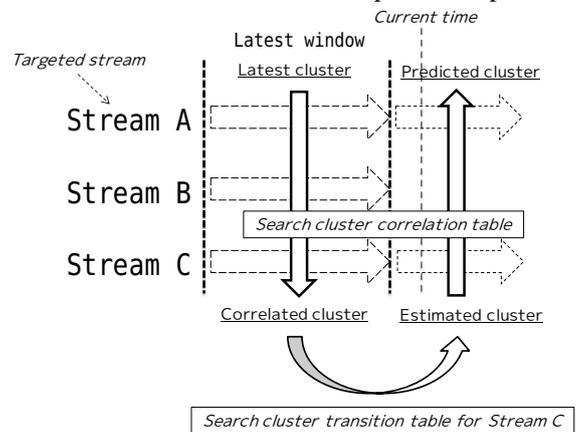
Figure 9: Stream Prediction by Correlation of Clusters
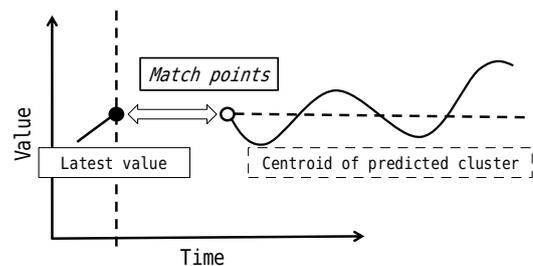
Figure 10: Restoration of sequences by Cluster Centroid

Table 1: Number of Clusters and Evaluation

Sensor	Scale	20 h		10 h		4 h		1 h	
		Num. of Clusters	Evaluation Value	Num. of Clusters	Evaluation Value	Num. of Clusters	Evaluation Value	Num. of Clusters	Evaluation Value
Temperature	0~40C	16	5.212	29	3.494	73	2.680	74	1.570
Humidity	0~100%	20	7.315	28	6.067	85	4.282	264	2.120
Illuminance	0~65000lx	20	1.819	43	1.945	**19**	**11.605**	**5**	**13.173**

Table 2: Average Number of Sequences in Each Cluster

Sensor	20 h	10 h	4 h	1 h
Temperature	3.062	3.370	3.342	12.568
Humidity	2.450	3.500	2.870	3.528
Illuminance	2.450	2.270	**12.842**	**196.000**

B) Partial Restoration of Sequences from Cluster

The method restores the next partial sequences from the predicted cluster according to the latest data and the cluster centroid of the predicted cluster. A cluster centroid is a specific partial sequence created from the average of all partial sequences in a cluster. This means the effective features of partial sequences in the cluster. Our method attaches the latest data to the cluster centroid, and regards the virtual partial sequence starting at the latest data as the next partial sequence, as shown in Figure 10. This enables the user and application to obtain the required data from the next partial sequence.

3.5 Approach to Multiple Interval Prediction

Feature quantity patterns express general features of partial sequences, but cannot determine the details. Thus, a cluster centroid cannot express the details of partial sequences associated with a cluster.

In addition, the predicted partial sequence length depends on the length of the cluster centroid used by prediction. Therefore, this depends on the length assumed by the classification system used in detection trends, because it is necessary to extract feature quantity patterns from partial sequences divided with the same interval as the interval of partial sequences to construct an accurate classification system.

According to this, the details on predicted partial sequences may differ from actual measured data. More specifically, this difference increases when the method predicts short partial sequences by using cluster correlation tables and cluster transition tables based on the classification system for long partial sequences. For example, the details on a centroid, which short partial sequences created, are lacking in long sequences, as shown in Figure 11.
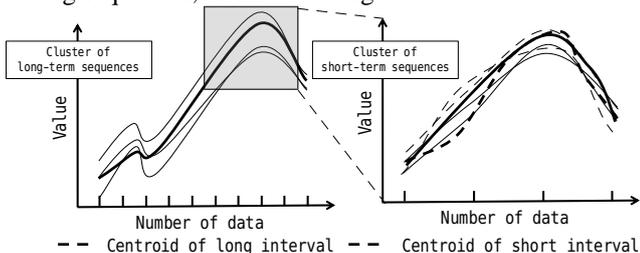
Figure 11:Data Lacking by Long-term Partial Sequences

Therefore, our method creates classification systems and tables with various partial sequence lengths. It selects the classifier and table according to the time when a user and application want to obtain data. For example, if data for the near future is required, the method uses the tables based on short intervals. This enables the user and application to obtain accurate predicted data.

3.6 Using Past Measurements

Our method detects correlations in data streams from cluster correlation table and cluster transition table. Because these tables require certain amount of past data for accurate detection of correlation, our method never detect correlations until past data are enough stored at these tables. Therefore, we use the clustering results that are used for building classifier to initialize two tables. We guess this attempt improve the accuracy of prediction when real time trend detection is launched.

If unique measurements which are cannot detect accurately by classifier appear, the accuracy of prediction will decrease. This case may often happen in data streams treating seasonal data, such as measurements of temperature sensor at natural environment. Therefore, we guess the classifier and two tables will be updated with a certain time.

4 EXPERIMENTS AND DISCUSSION

4.1 Evaluation of Clustering Algorithm

We carried out basic experiments to verify the efficacy of our method. We prepared sample data for the experiments and built classification system with various lengths.

This experiment used the stored data measured by a sensor unit used for farming [11] located in Fukushima Prefecture. This sensor unit measured the temperature, humidity, and illuminance from November 25, 2010 to February15, 2011. We only used 49 day of data from this span, when the sensor unit measured for 20 h between 0:00 to 20:00. The lengths of partial sequences were {20, 10, 4, 1}.

We defined the cluster evaluation value for prediction CE as the expectation value for prediction error obtained with Eqs. (3) and (4). Variable k is the number of clusters in these equations, and variable m is the number of partial

Table 3: Accuracy of Each Classifier

Sensor	20 h			10 h			4 h			1 h		
	J48	MP	BN	J48	MP	BN	J48	MP	BN	J48	MP	BN
Temperature	95.9	92.3	97.8	95.9	97.6	90.2	89.8	84.8	71.9	80.6	60.2	57.1
Humidity	100	100	90	97.9	97.6	92.0	92.9	90.2	85.5	90.0	49.7	65.6
Illuminance	97.5	93.9	90.2	100	92.3	77.0	87.2	84.2	77.4	55.1	76.1	76.0

sequences. In addition, variable c_{kj} means the j th data of the cluster centroid in cluster k, and x means one of the data. Variable x_{kij} means the j th data of the i th partial sequence at cluster k.

$$Dki = \left\{ \sum_{j=1}^{m} \frac{100}{(max - min)} \times \sqrt{(c_{kj} - x_{kij})^2} \right\} \Big/ m \quad (3)$$

$$CE_k = \left\{ \sum_{i=1}^{n} D_{ki} \right\} \Big/ n \quad (4)$$

Variable D_{ki} means the average of all deltas with all data of the cluster centroid. This variable is the delta of the i th partial sequence in cluster k. CE_k is the average of all deltas of partial sequences in cluster k. Variable max and min are max value and min are the maximum and minimum values according to sensor performance. A delta is expressed as a percentage.

Tables 1 and 2 summarize the results obtained from experiment. Table 1 lists the numbers of clusters and cluster evaluation values for prediction. Table 2 lists the expected number of partial sequences in each cluster. CE reached below 10 percent in most lengths of partial sequences. This means the error between actual data and predicted data will reach below at least 10 percent. For example, the error will reach bellow 4C in the temperature sensor. We assumed this error would be permissible for prediction in most cases. The expected numbers of partial sequences were near three in most lengths in Table 2. This is a particularly accurate result for our clustering algorithm that classifies the source into two clusters as recursively as possible. The numbers of clusters increase with decreasing interval lengths. This means unique types of partial sequences increase in inverse relation to decreasing interval lengths. Therefore, we attributed these results to our clustering algorithm being able to detect unique types.

In both tables, the illuminance sensor's results in cases where the interval length is 4 and 1 h differ from the other results. These accented results mean trend-detection failed in these cases. We assumed the reason our method missed is because the feature quantities used in it were not compatible with the illuminance data in these interval lengths. Therefore, the method must correctly select feature quantities according to the types of sensors and interval lengths.

4.2 Evaluation of Classifier Algorithm

We carried out a basic experiment on a classifier to verify an adequate algorithm for constructing it. Three classifiers were constructed in the experiment by using C4.5 and Multilayer Perceptron and Bayesian Networks, and we conducted 10-fold cross-validation to obtain the accuracy of the classifier. All algorithms were implemented in Weka. Multilayer Perceptron is an algorithm for building neural networks. Accuracy was expressed by the percentage of correct classifications. The experiment used the clustering results each interval of partial sequences.

Table 3summrizes the accuracy of each classifier where J48 is much more accurate than the others. Therefore, J48 was considered to be the best algorithm of all these. In addition, the best accuracy was over 90% at 20 and 10h. However, at 4 and 1 h, accuracy worsened. Most of the accuracy at 1 h especially fell below 80%. We assume the algorithms failed to construct classifiers in these cases. The reason the algorithms failed was because of our clustering algorithm, which could not take into consideration the threshold used by the classifier, because our clustering algorithm executed classification according to the Euclidean distance of the feature quantity pattern made from each partial sequence. We assumed the clustering algorithm would be ruled unsuitable as an algorithm for constructing classifiers by using a threshold. It is necessary to use the clustering algorithm by taking into consideration threshold of feature quantity to succeed in constructing classifiers.

5 CONCLUSION

This paper proposed a method of detecting the "timing correlation" between multiple data streams by using information on the appearance of trends. The basic experiments demonstrated the efficacy of the method of prediction using clusters created from the detection of trends. However, we found problems with the clustering method and feature quantities in clustering. It is necessary to do further research on adequate feature quantities for each sensor and intervals to find a new clustering method. In addition, we will verify the efficacy of our method by additional experiment. We will estimate cluster of past measured data by proposal method. The accuracy will be given by matching estimated clusters to actual clusters detected by classifier. If this experiment shows well-accuracy, we will implement the system predicts data stream in real time, which using our method, in order to verify the efficacy against single data stream prediction, where multiple data streams exist.

REFERENCES

[1] Y. Sakurai, "Stream mining technologies for time-series data," IPSJ Magazine, Vol47, No.7, pp.755-761 (2006). (*in Japanese*)

[2] H. Arimura, T. Kida, "Mining technologies for data stream," IPSJ Magazine, Vol.46, No.1, pp.4-11 (2005).

(*in Japanese*)

[3] Y. Fujiwara, Y. Sakurai and M. Yamamuro, "A search method for multiple data streams," Database Society of Japan Letters, Vol.4, No. 4, pp.13-16 (2006).

[4] K. Sugawara, Y. Shiraishi and O. Takahashi, "Estimation of the user action and static object using 3-Axis Acceleration sensor and vibration motor based on PWM control for cell-phone," Collection of papers of DPS Workshop 2009, pp.25-30 (2009). (*in Japanese*)

[5] H, Kawashima, M. Toyama and M. Imai, Y. Anzai, "Retrieval of similar sequences with waveform features," technical report of IECIE SIG-DE, Vol.102, No.209, pp.125-128 (2002). (*in Japanese*)

[6] M. Toyoda, Y. Sakurai and T. Ichikawa, "Stream matching based on Dynamic Programming," technical report of IPSJ SIG-DBS, Vol.2008, No.88, pp.277-28 (2008). (*in Japanese*)

[7] S. Papadimitriou, A. Brockwell and C. Faloutsos, "Adaptive, hands-off stream mining," Proceedings of the 29th International Conference on VLDB, pp.560-571 (2003).

[8] Y. Sakurai, S. Papadimitriou and C. Faloutsos, "BRAID: stream mining through group lag correlations," Proceedings of the 2005 ACM SIGMOD, pp.599-610 (2005).

[9] Y. Zhu, D. Shasha, "StatStream: statistical monitoring of thousands of data streams in real time," Proceedings of the 28th International Conference on VLDB, pp.358-369 (2002).

[10] S. Onishi, Y. Sakai and N. Yamaguchi, T. Shiraishi, "e-kakashi project," HTML, Available at "http://project.e-kakashi.com/" (2010) (*in Japanese*)3

[11] Machine Learning Group a University of Waikato, "Weka 3 - Data Mining with Open Source machine Learning Software in Java," HTML, Available at http://www.cs.waikato.ac.nz/ml/weka/

A Study on RPL-based Mutually Complementary Network

Naoki Yusa†, Hisashi Sawada‡, Kunihiro Yamada*, Tadanori Mizuno**, Hiroshi Mineno‡

†Graduate School of Science and Technology, Shizuoka University, Japan
‡Faculty of Informatics, Shizuoka University, Japan
* Professional Graduate School, Tokai University, Japan
** Faculty of Information Science, Aichi Institute of technology, Japan
yusa,sawada@minelab.jp, yamadaku@tokai.ac.jp, mizuno@mizulab.jp, mineno@inf.shizuoka.ac.jp

Abstract - We propose a RPL-based mutually complementary communication protocol (MCCP) for indoor sensor network under multi-interface communications. The MCCP reduces the constructing cost of sensor networks, and it improves communication reliability on the network. However, MCCP routing uses shortest hop route when initialize. On changing communication environment, it can't remake route. In this paper, we approached to apply RPL for routing protocol in lossy network to MCCP. Experimental results reveal RPL-based MCCP keep communication success ratio when environment changed

Keywords: Sensor network, MCCP, RPL, communication reliability

1 INTRODUCTION

During the recent years, sensor networks have been studied for various applications such as environmental monitoring [1], home network [2], healthcare [3], etc. Sensor network is a network consists of beaucoup sensor nodes, which can collect a measurement, observed real environment and personal information through a network. As a medium to construct a network, various things exist. But the sensor network using a lot of sensor nodes, the wireless communi-cation that the new installation of the communication infrastructure does not need are tends to be used as the sensor network using a lot of sensor nodes. However, wireless communication is weak in radio disturbance. And it is problem to increase the communication reliability of network. For this problem, there is Mutual Complementary Communication Protocol (MCCP) [4] using plural different communication mediums. It raises the rigidity for specific noise and redundancy of communication course and network topology by different communication mediums. The MCCP can increase communication reliability of the whole network. We approached to improve further communication reliability by using IPv6 Routing Protocol for Low power and Lossy Networks (RPL [5])-based MCCP.

2 RELATED WORKS

Several researchers have proposed routing protocol to communicate in unstable environment. These protocols on the premise of using communication interfaces such as ZigBee, IEEE802.15.4, and PLC. These interfaces are not stable.

2.1 MCCP

MCCP is network protocol to use different types of communication mediums complimentarily mutually. In MCCP, It assumes to use wired communication and wireless communication as a communication medium. It assumes in specialty the use of the power line communication as wired communication. In a sensor network assuming a lot of sensor node participated, wireless network that no limit to place and low cost to construct network infrastructure is necessary. However to use only wireless network to communicate, there are problems difficult to resolve like exogenous noise, interference by each communication, blocking radio wave by wall. Therefore maintenance to keep up communication reliability as the network is big problem. Using wired communication like Ethernet, optical fiber communication to sensor network, is robust to exogenous noises and no interrupted by wall and roof because they were covered by cable for exclusive use. Wired communication is able to communicate where wireless communication cannot. But wired communication need to construct infrastructure. It increases cost to use sensor network. If using Power Line Communication (PLC) as wired communication, network can construct with few cost because it uses already constructed power line. The other hand, PLC is interrupted by electric appliance connecting power line and cannot communicate different phase. Therefore maintenance to keep up communication reliability same as the network is big problem as wireless network.

2.1.1 MCCP network

MCCP constructed by two logical devices called end device and router device. Between router device and end devices ware made up to star topology, between routers ware made up to mesh topology. Router device's works are administration network and to relay packets on network. The other hand, end device only communicate messages with router device administrating own subnet. The MCCP assumes router devices have both wired communication interface and wireless communication interface, end device have either wired communication interface or wireless communication interface. But all devices must not have these interfaces as MCCP process. End devices are not required to have large memory space and high processing unit for network administration and to relay packets. In addition, device works only for communicate own message. Therefore, Device can cyclic sleep and work for long term with battery power.
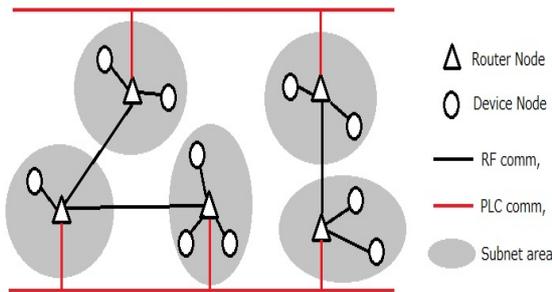
Figure 1: Network topology overview

2.1.2 Address conformation

The MCCP has another address to use in MCCP in addition physical address used in communication interface. MCCP address is constructed by Subnet ID and Device ID. Subnet ID is allocated each subnet in MCCP network. Device ID is allocated to participated End devices in each subnet. Subnet ID will allocate unique ID for each subnet. Device ID is allocated redundant ID will allocate. Above MCCP network layer, physical address will be shielded and only MCCP address will enable to use to communicate.

2.1.3 Merit and Weak point in MCCP

To use MCCP to sensor network obtain following benefits:

- Low cost sensor network. Both wireless communication and wired communication have to construct new infrastructure.

- Scalable network. In case of wireless communication cannot communicate, wired communication compensates. Wireless communication compensates communication between different phases of PLC.

- Increase communication reliability. Wireless and wired communication are not stable communication way. But, decreasing of communication quality can compensated by another communication, because each obstacle is different.

- Long network lifetime. Devices using PLC is connected power line and can work anytime. Network administrated by these devices, decrease load of battery powered devices and extend system sleep time.

And the other hand, MCCP weaks to change communication environment. Device joins to MCCP network by using Discover, Advertise, Join, Offer and Confirm messages. The MCCP network cut off when communication environment changing to any interface cannot communicate. It dues to MCCP has no way to redeem dead communication path.

For this problem, we proposed to compensate network by using RPL and changing communication route dynamically by environment.

2.2 RPL

RPL is Routing Protocol for Low power and Lossy Networks based on IPv6. ROLL [6] working groups on IETF is plotting it now. RPL use Destination Oriented DAG (DODAG), a DAG (Directed Acyclic Graph) rooted at a single destination, as network topology. Figure 2 shows DODAG.
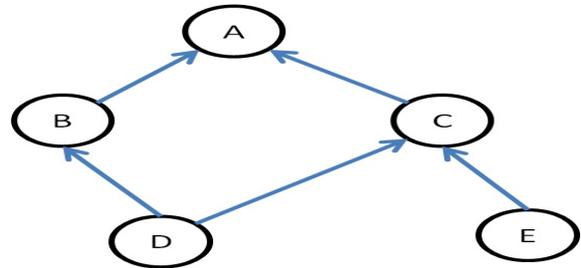
Figure 2: Example of constructed DODAG

Communication routes selected by each node's Rank on DODAG. Rank can calculate with a constellation of metrics. Moreover, it doesn't limit to metric, and RANK can be decided by constraint the device. The residual quantity of the battery etc will be enumerated as constraint. It can construct network adequate for the purpose, for using weighting factors to these parameter when Rank decided. For example, use parameter as data sending frequency, decrease unevenness of trafic and can use network resource efficient. The calculating formula of Rank is shown below.

$$\text{Rank}_{own} = min(Rank_{parents}) + \sum Param_i \times \alpha_i$$

Param:Metrics or constraints
$\alpha : weight\ factors$

Each node of RPL holds neighbor information in DODAG. Nodes having lower rank than own one are called parent nodes. And the lowest rank node is called preferred parent and used to communication. In addition to parent node, nodes having higher rank are called child nodes. Figure 3 shows example of DODAG and information of each node. The function group of the calculating formula etc. of RANK is called ObjectiveFunction (OF), and common ID (ocp: Objective Code Point) is allocated by the content of the processing of OF. Moreever, child'sANK must higher than own one, and parent's Rank must lower than own one. And there have to be a DODAG root allocated the lowest rank in DODAG. When Loop is generated in DODAG, Loop can be detected by collapsing this condition. In that case, Loop can be resolved by resetting held information, and recalculation of RANK. In addition to these loop protections, DODAG root node can restructure current own all DODAG. Each RPL node holds information of own parent and child nodes and route to child node. Except, the RPL root node has no parent node. Route information has child node and transfer node for sending to child node. Some nodes like terminal node in DODAG have no route information.
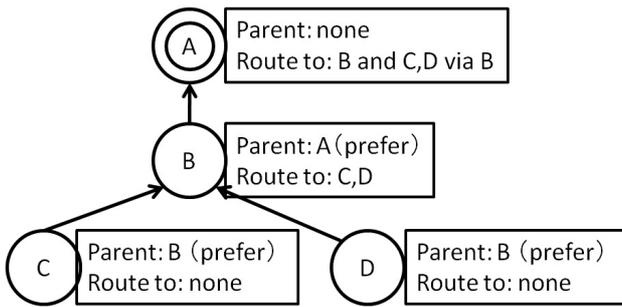
Figure 3: Example of DODAG and information of each node.

Figure 4: Packet format of RPL-based MCCP.

Rule of sending packet is looking up for destination node in own neighbor node information at first. If node found, send packet directly. If not, looking up for own routing information to destination node. If route exist, send packet in accordance with route information. In case of destination not found on holding information, transit packet to own preferred parent. And parent node receiving transmitted packet compare destination address and own one. If two addresses agreed, packet will be received. If address does not agree, parent node does same sending process. By this iteration, packet will be translated to destination node finally.

in this paper, construct DODAG over MCCP, and communicate by it.

3 APPLYING RPL TO MCCP

We approached compensating network by applying RPL to the MCCP and constructing DODAG over MCCP. We applied RPL to MCCP based on sample and draft of open to the public. There are big differences between sample and MCCP as follow:

- RTOS for device. The MCCP uses TRON, but RPL sample uses Contiki OS as RTOS. Therefore, we added callback timer function like Contiki OS. This function can use as function that Contiki OS has. And this timer can add while being free memory space.

- Network protocol. Between MCCP and sample of RPL has some differences originating by protocol.

- Construction of address. MCCP use Subnet ID and Device ID as address. But RPL use IPv6 address. To use sample effectively, we used fake IPv6 address made from Subnet ID and Device ID of MCCP.

- To detect invalid parent. RPL uses NUD (Neighbor Unreachability Detection) [7] to detect invalid neighbor node. But, the MCCP has no method like NUD. Therefore, we gave timer to each parent and parent will be expired when timer runs out.

3.1 Packet Format of RPL-based MCCP

RPL leverages DIO (DODAG Information Object), DIS (DODAG Information Solicitation), DAO (Destination Advertisement Object) messages for structuring and maintenance DODAG.
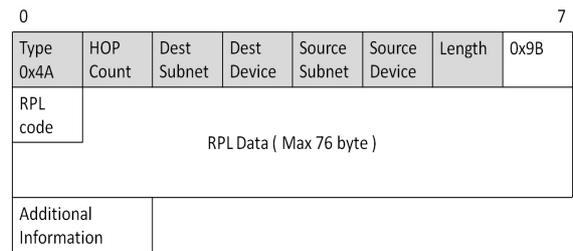
DIO message includes sender's DODAG information, rank and so on, and RPL node uses this information to join DODAG. This message will be sent periodically by multi casting. The node receiving these messages, decides own preferred parent and own Rank by DIO information, and sends DIO message soon by multi casting. Interval of sending DIO message can be longer in stable network. And decrease packet traffic for maintenance DODAG. Interval of sending DIO message will elongated by sending DIO message and problem happened in DODAG, it will reduced.

DIS message solicits to send DIO message. The node receiving these messages, send DIO message soon regardless of interval. The new node can join in DODAG quickly by using this message.

DAO message pass on own information from child to parent node. This message will send to own preferred parent by uni-cast. The node receiving these messages transmits messages to preferred parent. And add route information, as destination node, the node published DAO message and as via node, the node transmits this message.

We added MCCP packet type to distinguish these RPL affiliated message. Figure 4 shows the packet format of RPL-based MCCP. Gray colored area means MCCP format. RPL code in the packet is used to specify the message content. It is shown Table 1.

Table 1: Association of RPL code and message content

RPL code	RPL Message content
0x00	DIS
0x01	DIO
0x02	DAO
0x03	DAO ACKnowledgement
0xEE	MSG(User data)

We created MSG packet to transmit message by RPL routing rule demonstratively. Figure 5 shows format of MSG Packet. Target Subnet and Target Device mean destination of this packet. And these are different from Dest Subnet ID and Dest Device ID, and so on in the MCCP format. The MCCP' one means node information of sending this packet and will be altered by transmitting. Message area will be used for user data.
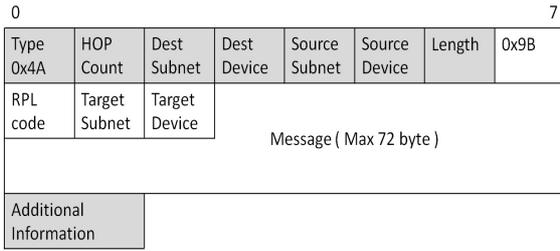
Figure 5: Format of MSG packet on RPL-based MCCP

3.2 Setting of Lifetime

In this verification of RPL-based MCCP, we use Table 2. These units are second. Our setting values are short as compare to sample, because we change communication environment as short interval.

Table 2: Setting for this verification

Setting Item	RPL-based MCCP	sample
Route info lifetime	60	NUD
Parent info lifetime	120	NUD
Initial interval to send DIO packet	$32 \sim 64$	$1024 \sim 2048$

4 SIMULATION

4.1 Development Module

We choice EM9 [8] for verification of RPL-Based MCCP.

The EM9 is emulator complying with routing protocol software. This consists of SubServe and Rtwin and works on Windows PC. SubServe emulates for communication between nodes. And Rtwin emulates MCCP device operation. These processes communicate by using API. Rtwin designed to use program code of real MCCP device. Therefore, EM9 can emulate the entire MCCP operation by using these. Figure 6 shows whole of EM9.

However, it has some restrictions. For example, functions depended on hardware cannot use. And communication between Rtwin and real machine is impossible. In this verification, these functions will be never used, we choice the EM9. We prepared following rules of environment for verification.

- Communication environment of initial state shown in Figure 7

- Communication environment will change in every 10 minutes. A state of unstable way will be flipped. when there is only one enable route in unstable routes, another unstable route will be changed. Therefore, all unstable way will not be disabled in same time. A metric of flipped way will change in EM9.

- We use Signal Quality as metric. But in EM9, these metric data has no effect on result of communication.
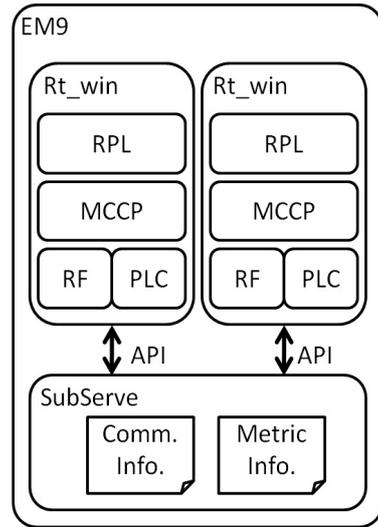
Figure 6: Whole of EM9
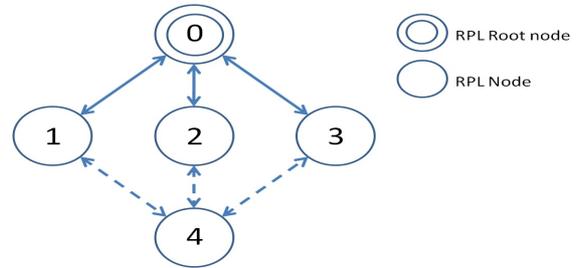
Figure 7: Initial communication state. Number in circle shows Subnet ID of node. Arrow line shows communicable between two nodes. And it means stable communication route. Dash arrow line shows communicable too. But it means unstable communication route. When communication environment changed by emulator, these are candidate to be flipped.

Therefore even if signal quality is very low, they can communicate. It will used for calculate rank.

$$\text{Rank} = \min(\text{Rank}_{parents}) + SignalQuality \times 0.1$$

- Send MSG packet every random interval (from 10 to 20 second). Message area includes sequence number in MSG packet. Node 4 will send this to Node 0. Node received MSG packet, will log to file with time and message, publisher information.

- Use only RF as communication interface.

Simulation Result We verified to be constructed DODAG topology at first and it was constructed. Figure 8 shows sequence of RPL message while constructing first DODAG. After these sequences and constructing DODAG, each node's rank will be decided by metric information sent by DAO packet. Moreover, metric information in DAO message includes RF and PLC's one. But we use only RF's one in this verification.

Figure 9 shows actual constructed DODAG topology and Table 3 shows information in each node after 5 minutes of beginning. PP means Preferred Parent. NBR means neighbor node
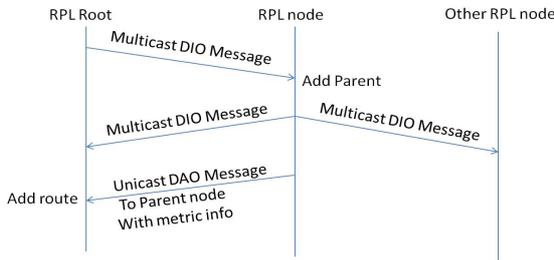
Figure 8: Sequence of RPL message
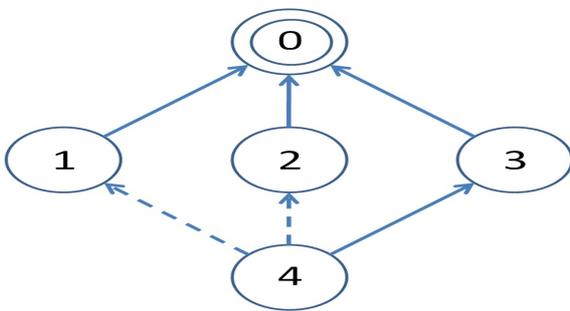
Figure 9: Constructed DODAG topology

Table 3: Information in each node

Number of node	Rank	Parents info.	route info.
0 (RPL root)	100	none	To 1,2,3(NBR) To 4 via 3
1	157	0(PP)	To 4(NBR)
2	144	0(PP)	To 4(NBR)
3	106	0(PP)	To 4(NBR)
4	144	1,2 3(PP)	None

We verified for four hour and counted unreached packets from sequence number in message. Figure 10 shows result of verification.

In this verification, MSG packets were sent 855 times and total unreachable packets were 84. Success rate of communication was 90.1 percent when calculate from these data. These results show RPL-based MCCP can retain communication way when communication environment changed.

5 CONCLUSIONS

In the MCCP of the communication protocol, for the indoor sensor network, we approached to improve of the communication reliability by using RPL. The RPL composes DODAG
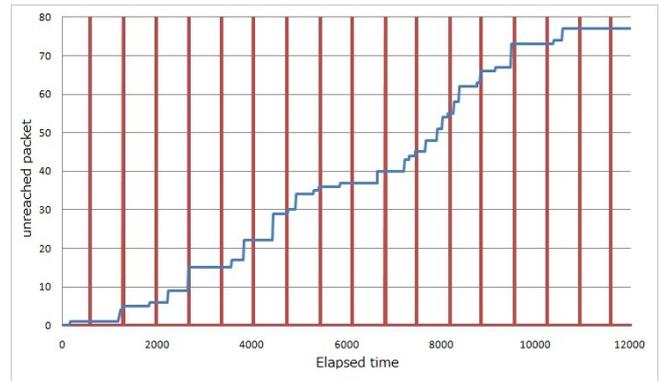
Figure 10: Elapsed time in second vs. accumulated total of unreached packet. Vertical red line indicates to flipped communication environment.

topology and communicates based on it. Moreover, to construct and maintenance, the RPL message is used. We thought to homologize by using these RPL functions on MCCP not assuming the changing of communication environment. Then we applied RPL to MCCP and verified with changing communication environment.

As a result, the construction of DODAG on the MCCP network was succeeding. And about 90 percent of message were reaching to destination. These results show RPL-based MCCP can keep communicating even if communication environment changed quickly. And it can increase communication reliability. We think some strategies to improvement of RPL-based MCCP in this paper. In this verification, we selected only RF interface. Communication reliability will be able to increase much further by using both RF and PLC interface [9]. In that case, we think to need some idea to calculate rank, for example to use better metric than other one. In addition to, each lifetime of parent and route were fixed values. By changing dynamically these parameters by variation of metric information, each node will select more stable route and communication reliability will increase.

REFERENCES

[1] G. Barrenetxea, F. Ingelrest, G. Schaefer,M. Vetterli, O. Couach, andM. Parlange, "SensorScope: Out-of-the-Box Environmental Monitoring," in Proc. of IPSN '08,pp. 332- 343, Apr. 2008.

[2] D. Lymberopoulos, A. Bamis, and A. Savvides, "Extracting Spatiotemporal Human Activity Patterns in Assisted Living using a Home Sensor Network," in Proc. of the International Conference on Pervasive Technologies Related to Assistive Environments (PETRA), 2008.

[3] S. Jiang, Y. Cao, S. Iyengar, P. Kuryloski, R. Jafari, Y. Xue, R. Bajcsy, and S. Wickler, "CareNet: an integrated wireless sensor networking environment for remote healthcare", in Proc. of ACM BodyNets '08, Mar. 2008.

[4] Kuriyama, H., Mineno, H. and Mizuno, T., "Evaluation of Mutually Complemen-tary Multichannel Sensor Network forWireless and Power lines", in proc. of IEEE In-

ternationalSymposium on Power Line Communications and its Application(ISPLC),pp.223-227 2009

[5] Routing Over Low power and Lossy networks, "draft-ietf-roll-rpl-19", in IETF, 13, Mar 2011

[6] ROLL working group http://tools.ietf.org/wg/roll/

[7] RFC2461, Neighbor Discovery for IP Version 6 (IPv6), December 1998

[8] EM9, Emulation software for routing protocol software, Renesas Technology Corp. and Renesas Solutions Corp.

[9] Hisashi Sawada, Hiroshi Kuriyamay, Yuki Nakanoz, Hiroshi Minenoz and Tadanori Mizunoy, "A Mutually Complementary Communication Protocol for Indoor Sensor/Actuator Networks", in The 11th IEEE /IPSJ International Symposium on Applications and the Internet 18-21, July. 2011

Session 4:
Computer Assisted Work
(Chair: Yoh Shiraishi)

LASP — a Learning Assistant System for Formal Proof

Kiyoyuki MIYAZAWA[†], Kozo OKANO[†], and Shinji KUSUMOTO[†]

[†]Graduate School of Information Science and Technology, Osaka University
Yamadaoka 1–5, Suita City, Osaka, 565-0871, Japan

Abstract - The basis of formal techniques is mathematical logic. Especially, it is important to understand the concept of the formal proof. It is, however, difficult for novices to study formal proof because of rigorous operations. To resolve such a problem, we propose and implement a prototype of a learning assistant tool for formal proof, called LASP. The purpose of LASP is to prevent the drawbacks which occur when normal learners exercise by the hand. The advantages of the proposed method are the following: 1. input support for long logical expressions; 2. the users are not required to operate copy and paste when they make proofs; 3. hint features facilitate constructing proofs; and 4. the proofs are able to output in LaTeX files. Experimental results show that LASP improves the drawbacks of conventional exercises.

Keywords: Logics, Formal Proof, Computer Aided Education

1 Introduction

Formal approach is a core technology of modern software development. There are mainly two approaches as the formal approach technique: interactive theorem proof and model based approach such as model checking[2]. Coq[6], Agda[7], Isable/HOL, and so on, are famous interactive theorem proof systems. They require understanding formal proof of logic, of users. The model checking also requires logic of user in order to represent properties on the target system. Other model based approaches also need logics to describe constraints on models representing the target systems and software.

On the other hand, it is said that "Mathematical reasoning is intrinsic to both traditional engineering and software engineering, . . . Software engineers usually use discrete mathematics and logic in a declarative mode for specifying and verifying system behaviours and for analysing system features"[11]. This illustrates that it is important to understand mathematical logic in software engineering. Supporting the learning of mathematical logic can promote the learning of basic techniques of the formal approach. We can promote productivity of software by learning the formal approach.

A learning assistant system which we developed in this research supports to learn formal proof based on Hilbert's axiom system. As related works, Jon Barwise et al. have developed Turing's World[4] and Tarski's World[3], [5]. We can learn graphically Turing machine using Turing's World. Also we can learn graphically semantics of mathematical logic using Tarski's World.

Tarski's World shows simple 3D computer graphic worlds where geometric blocks of various kinds and sizes are allocated. Tarski's World gives you a first-order predicate sentence with a 3D view and let you decide whether they are true or false in the view. For example, a sentence "there exists a cube among the all objects" is given, a user has to decide the sentence is true or not by seeing the view. Three-dimensional views are sometimes used in Computer-Aided Education. For example, paper[12] provides a 3D geometric construction tool specifically designed for mathematics and geometry education. It is based on a mobile augmented reality (AR) system.

Another approach for computer-aided education, is using collaborate web-based environment, such as Moodle and WEB-CT. Several other approcahes are proposed. For exmple, paper [10] provides a flexible learning scheme for selected pilot courses in engineering education. In such a scheme, traditional lectures and written exercises are combined with additional Web-based learning resources.

As for logic orient education, MacLogic[9] supports the learning of natural deduction of Gentzen. Logic tutor [1] supports students to learn logic. It has a feature to analyse and restore a student's mistakes during solving logic exercises. It has been used in a logic class in Department of Computer Science, University of Sydney.

We target novices of logic and provide functions for them. Therefore, we don't think that the idea applying simply tools like conventional theorem provers [6], [7] to education is suitable, although there is a similar approach. For example, "Bringing Research Tools into the Classroom"[14] helps moving computational tools used for research into the classroom. It successfully bring high performance computing into modelling courses, built software for both protein structure visualization and hydrological analysis of watersheds and so on. Project[8] aims at supporting the entire process of mathematical theory exploration within one coherent logic and software system. It uses formal logic and computer aided approach; it helps a learner to understand core of mathematics.

While related works mainly focus on understanding semantics, we focus on understanding formal proof. For understanding semantics of logic, the approach like Tarski's world, which utilizes 3D graphic is useful. We, however, focus on formal proof; 3D graphics might help a little. In addition, our tool deal with Hilbert's axiom system, because the class of mathematical logic held in Department of Informatics and Mathematical Science, Osaka University deals with Hilbert's axiom system only. We could not find any tools to learn Hilbert's formal proof as far as we know. Therefore, we developed our tool and obtained evaluation.

The paper is organised as follows. Section 2 presents a brief introduction of formal proof. Section 3 and 4 provide our tool and experimental evaluation. Section 5 presents discussion. Finally Section 6 concludes our paper.

Table 1: Axioms for FOP

A1	$P \rightarrow (Q \rightarrow P)$
A2	$(P \rightarrow (Q \rightarrow R)) \rightarrow ((P \rightarrow Q) \rightarrow (P \rightarrow R))$
A3	$(\neg P \rightarrow \neg Q) \rightarrow ((\neg P \rightarrow Q) \rightarrow P)$
A4	$\forall x\, T(x) \rightarrow T(t)$,
	where term t is free from variable x of $T(x)$.
B1	$P, P \rightarrow Q \vdash Q$ \qquad (MP, modus ponens)
B2	$P \rightarrow Q \vdash P \rightarrow \forall x\, Q$,
	where term P is free from a free variable x

2 Formal Proof

The formal proof is a process which proves theorems by axioms and inference rules. The axiom is a major premise to derive a concrete logical expression instance without contradictions; while the inference rule is used to derive a new logical expression from proved ones. For example, modus ponens is an inference rule which derives Q from $P \rightarrow Q$ and P (P and Q are arbitrary logical expressions).

An axiom system consists from axioms and inference rules and the logical expressions proved by an axiom system is called the theorem. Figure 1 shows an example of formal proof which proves a theorem of $\vdash X \rightarrow X$.

Table 1 shows the axioms and inference rules for first order logic by Hilbert.

3 LASP a Learning Assistant System for Formal Proof

In this section, we will present our tool "Learning Assistant System for Proof, (LASP)."

3.1 System Overview

We developed LASP to reduce problem practice time by handwriting. It is based on Hilbert's axiom system.

The defects of practice of formal proof by handwriting as following.

1. We may make a mistake in writing of long expression corresponding to an axiom, especially correspondence between terms and variables.

2. We are often required operations like copy and paste when we construct a formal proof.

Resolving the problems above, we define the goal as follows.

1. LASP has an interface to minimise inputs of logical expressions.

2. We are not required operations like copy and paste.

LASP is implemented in Java with 8660 LOC. The system is published in WEB[13].

We will describe those features and interfaces specifically.
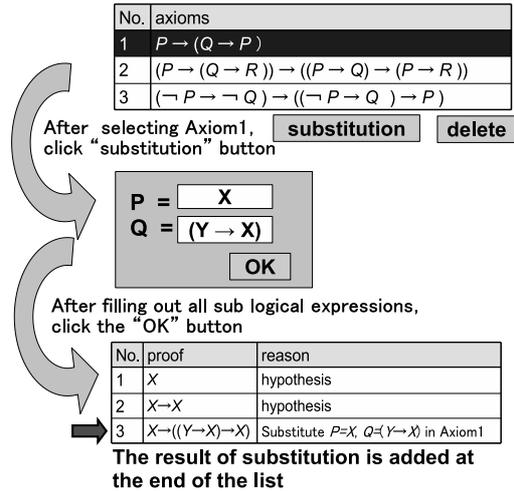
Figure 1: Substitution

3.2 Substitution Assistant

We implemented a substitution assistant feature in order to reduce useless proof time which we devote to avoid careless written mistakes such as inconsistency of corresponding of sentence structure or variable name. Such mistakes also often occurs in handwriting complex and long logical expressions. These are unrelated to learning essential proof. Instead, it is important to learn thinking process of problem resolution in mathematical logic.

Let us consider the following example. Substitution of variables P, Q and R of axiom A2 in Table 1 with term $X \rightarrow Y$, Z and $((Y \rightarrow Z) \rightarrow X)$, respectively, yields a logical expression (1), which is obviously long and complex. It also hard to read and write.

$$
\begin{aligned}
((X \rightarrow Y) \quad &\rightarrow \quad (Z \rightarrow ((Y \rightarrow Z) \rightarrow X))) \\
&\rightarrow \quad (((X \rightarrow Y) \rightarrow Z) \rightarrow ((X \rightarrow Y) \\
&\rightarrow \quad ((Y \rightarrow Z) \rightarrow X)))
\end{aligned} \tag{1}
$$

Substitution Assistant automatically generates new logical expressions when a user selects an arbitrary axiom and inputs arbitrary logical expressions as substitution terms for the selected axiom.

Figure 1 shows a substitution flow on LASP.

Axioms are managed using a table as in the upper of Figure 1. When a user clicks an axiom which he wants to use from the table, then Substitute Panel opens. If he inputs logical expressions into every propositional variable on Substitute Panel and clicks OK button, then the result expression is added at the end of the proof list.

Such a feature has been already built in many interactive theorem proof systems. They, however, are not easy to use in under-graduate classes.

3.3 Support for First-order Logic

LASP supports not only propositional logic but also first-order logic.
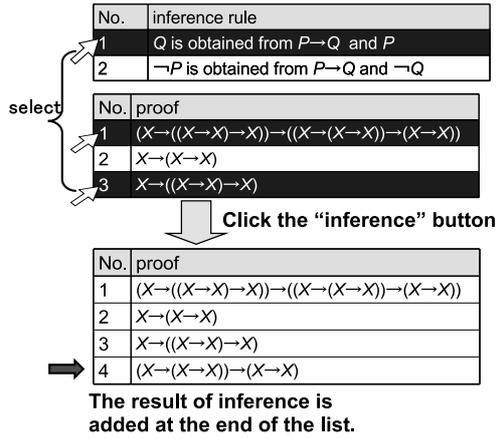
Figure 2: Flow of constructing a proof with LASP

We have to consider the variable is free or bound when we want to perform substitution for a variable of an expression in first-order logic. For example, variable y of a logical expression (2) is bound by $\exists y$. Therefore, this variable cannot be substituted.

$$\forall x \exists y f(y, z) \qquad (2)$$

Furthermore, any term included variable x or y cannot be also substituted in variable z of an expression (2) because such substitution resulted bound by the quantifier. If expression $g(x)$ was substituted into z, the result would be a logical expression (3).

$$\forall x \exists y f(y, g(x)) \qquad (3)$$

In the expression variable x is bound by $\forall x$ and semantic of the expression (3) is different from the expression (2). LASP generates an exception if an illegal substitution like above is occurred by using substitution inhibition list.

3.4 Inference Assistant

Inference Assistant is implemented to reduce practice time by reducing operations like copy and paste in handwriting.

When a user selects inference rule that he wants to apply and proven logical expressions that conform with inference rule, a new logical expression that corresponds to its rule is generated. The reason of generating is also generated and it reduces time to handwrite reasons.

Figure 2 shows the flow of the application of inference rule on LASP. Inference rules and proven theorems are displayed in a table. First, a user clicks and selects an inference rule which he wants to apply from the inference rule table. Second, he clicks and selects proven theorems matching inference rule. Finally, he clicks inference button. If he selects proven theorems rightly, then a new logical expression is added at the end of the proven theorem table.

Figure 3 shows the screen-shot of LASP.

3.5 Deduction Theorem Assistant

Deduction Theorem is a useful theorem to prove a new theorem efficiently. Therefore, LASP also supports Deduction
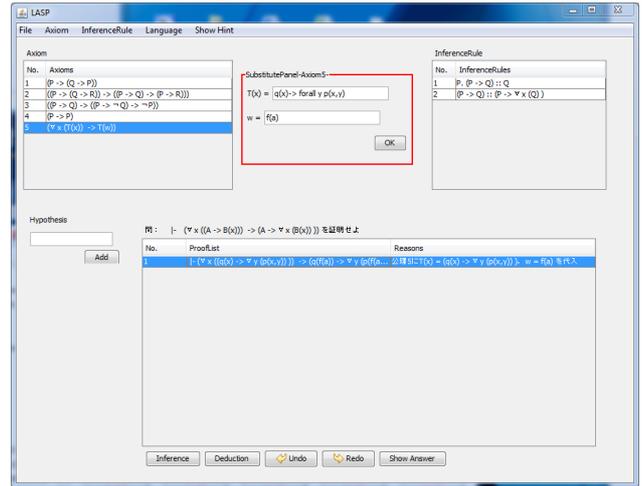
Figure 3: Screenshot of LASP

Theorem. The following is Deduction theorem.

From the fact that $\Gamma, P \vdash Q$, we obtain $\Gamma \vdash P \to Q$.

First, a user selects a proven theorem that he wants to apply to Deduction theorem. Second, he clicks the deduction button. Then, deduction panel opens. He checks the radio buttons to select theorems applying Deduction Theorem and clicks the OK button. Finally, the result of applying to deduction theorem is added at the end of the proven theorem table.

3.6 Hint Features

These features provide three levels of hints for users who is unfamiliar with formal proof. They are implemented after obtaining a feedback from Experiment 1 described at section 4. The feedback is that it is difficult to solve the exercises which have many steps by oneself. The hints prepared by LASP is as follows.

1. fill-in-the-blank hint

2. milestone hint

3. providing next hint only

Table 2 shows the fill-in-the-blank hint for a proof of $A \to \neg\neg A$ under the assumption $P \to P$ has already proven.

It is generated randomly when LASP reads data of exercises.

Table 3 shows the milestone hint for the same proof. The hint shows expressions that will be proven by applying inference rule.

We assume that a theorem $P \to P$ is given as axiom A4.

Table 4 shows a typical solution.

3.7 Undo/Redo Features

LASP supports undo/redo features. Learners can revise their proofs using these features.

It would be better that LASP supported a feature in which a user can edit his/her proof after he/she have proven. However, this feature might distory reasoning chain of the proof.

Table 2: A hint of fill-in-the-blanks

1		Hypothesis
2		"?" is substituted in axiom "?"
3		"A" and "?" are substituted for "P" and "Q" each other in axiom "1"
4	$(\neg A \to A)$	Applying the inference rule "?"
5		"?" and "?" substituted for "?" and "?" in axiom "?"
6	$((\neg A \to \neg A) \to \neg\neg A)$	Applying the inference rule "?"
7		Applying the inference rule "?"
8	$(A \to \neg\neg A)$	Deducing "?" from proven theorem "?"

Table 3: A hint of milestone

1		
2		
3		
4	$(\neg A \to> A)$	Applying the inference rule "1" to proven theorem "1" and "3"
5		
6	$((\neg A \to \neg A) \to \neg\neg A)$	Applying the inference rule "1" to proven theorem "4" and "5"
7	$\neg\neg A$	Applying the inference rule "1" to proven theorem "2" and "6"
8		

Thus, the current LASP only support Undo/Redo features. Undo/Redo works in as many steps as the length of proof users made.

3.8 Input/Output Features

Logical expressions are often long and deeply nested. Therefore, string expression of a logical expression is sometimes difficult to understand. This feature visualises their parse trees (Fig. 4). Clicking a node, the sub tree is expanded or shrink (Fig. 5).

LASP also support LATEX format for output constructed proof. It also helps users, especially describing reports on homework.

The exercises of LASP are input as XML files. Therefore, a teacher easily prepares new sets of exercises with hints.

4 Evaluation

This section describes the evaluation. We have performed two experiments. Experiment 1 is conducted in a class of mathematical logic. Subjects are 50 undergraduate students. Experiment 2 is conducted after reflecting the feedbacks of Experiment 1 and the number of subjects are 16. The subjects
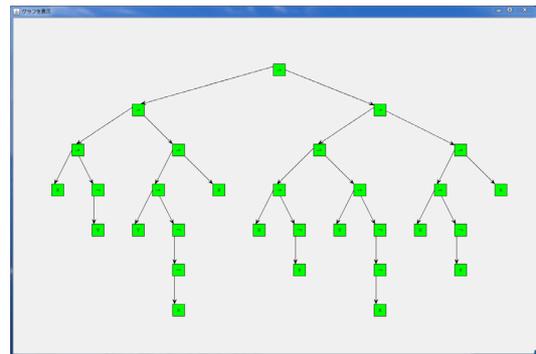
Figure 4: Parse Tree Representation
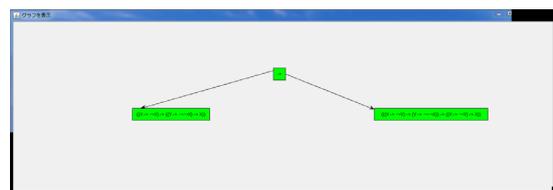
Figure 5: Parse Tree Representation (Shrinked Tree)

Table 4: The Solution

1	A	Hypothesis
2	$\neg A \to \neg A$	P is substituted in axiom A4
3	$A \to (\neg A \to A)$	"A" and $\neg A$ are substituted for "P" and "Q" each other in axiom "1"
4	$(\neg A \to A)$	Applying the inference rule "1" to proven theorem "1" and "3"
5	$((\neg A \to A) \to (\neg A \to \neg A) \to \neg\neg A)$	"$\neg A$" and A are substituted for "P" and "Q" each other in axiom "3"
6	$((\neg A \to \neg A) \to \neg\neg A)$	Applying the inference rule "1" to proven theorem "4" and "5"
7	$\neg\neg A$	Applying the inference rule "1" to proven theorem "2" and "6"
8	$A \to \neg\neg A$	Applying Deduction Theorem

in Experiment 2 are a doctoral student, ten master course student and four undergraduate students.

4.1 Goals of the Evaluation

The objective in these experiments is as follows: to measure the degree of effectiveness for users, and to collect feedbacks to make LASP more useful.

4.2 Items of the Evaluation

In Experiment 1, we mainly investigated how effectively a user can solve an exercise and the usability of LASP. The degree of efficiency is measured by the time that subjects finish solving given problems.

As the evaluation of usability, we prepare questionnaires to research how subjects can be familiar with LASP and what kind of interface is needed to enhance usability. The items of the questionnaires are summarised as follows.

$Q1$ The degree of user-friendliness of Substitute Assist.

$Q2$ The degree of user-friendliness of Inference Assist.

$Q3$ The degree of user-friendliness of whole tool.

$Q4$ The degree of efficiency by using LASP.

$Q5$ The degree of reducing trouble in proving by using LASP.

$Q6$ The degree of effects of learning formal proof.

The items are on a scale of one to five. Five means the strong degree. A free comment space to collect opinions about LASP or the experiment is also included.

4.3 Procedure of Experiments

The procedure of Experiment 1 are as follows.

1. We divide all subjects into six groups.

2. Subjects solve two problem practices by the hand.

3. Two weeks later, subjects solve two problem practices by LASP.

   (a) We distribute tool manual to all subjects.

   (b) We let subjects use LASP by solving sample exercises for 20 minutes.

   (c) Subjects solve the exercises.

4. Subjects answer the questionnaire.

There are four exercises in all. The order of exercises is randomly chosen to each group. Time limit that a subject can solve an exercise is 15 minutes.

The procedure of Experiment 2 is as follows.

1. We distribute tool manuals and textbooks about mathematical logic to all subjects.

2. We let all subjects solve four sample exercises for an hour and familiarise them with LASP.

Table 5: The Results of the Questionnaires in Experiment 1 and 2

item	Exp.1	Exp.2
$Q1$	3.70	4.69
$Q2$	3.28	4.56
$Q3$	2.84	4.13
$Q4$	2.88	3.63
$Q5$	3.28	3.75
$Q6$	2.84	3.63

Table 6: The problems in Experiment 1

	problem
$Q1$	$A \rightarrow \neg\neg A$
$Q2$	$A \rightarrow B \vdash \neg B \rightarrow \neg A$
$Q4$	$\forall x(P(x) \rightarrow Q(x)) \rightarrow (\forall x P(x) \rightarrow \forall x Q(x))$

3. Subjects solve exercises by the hand and using LASP alternately.

4. Subjects answer the questionnaires.

The time to familiarise subjects with LASP is an hour in Experiment 2 because twenty minutes were not enough to familiarise in Experiment 1. There are four exercises in all as well as Experiment 1. The order of exercises is different respectively by group. The way of measuring solving time is the same as Experiment 1.

4.4 Results of the Experiments

4.4.1 Results of the Questionnaires

Table 5 shows the average points of items of questionnaires in Experiment 1 and 2.

The following feedbacks are obtained from the free comment space of Experiment 1.

- It is better that nest of logical expressions shows clearly by colour-coded font.
- It is troublesome that users have to use mouse.
- Users want to adjust window size freely.
- Panel alignment should be improved.
- Shortcut keys should be provided.
- Hint feature should be provided.
- It must show an answer.
- It seems to be able to be familiarised users by taking enough time.

The following feedbacks are obtained from the free comment space of Experiment 2.

- LASP should notify clearly when the proof is correctly constructed.
- It is troublesome for users to write expressions including quantifiers because users have to input "forall" to display $\forall$.
- Users are not allowed to delete logical expressions except the latest one.

Table 7: The number of subjects who solve correctly in Experiment 1

	handwriting	LASP
$Q1$	5/23	7/23
$Q2$	6/24	7/23
$Q4$	0/27	1/22

Table 8: The average time of solving in Experiment 1

	handwriting	LASP
$Q1$	9m13s	10m04s
$Q2$	6m56s	8m15s
$Q4$	Nobody solves	8m18s

- Users are not allowed to add logical expressions except at the end of one's list.
- It is better that there is a memorandum panel to think proof tactics.
- It is better that there is a feature that user can make his own exercises.
- Shortcut keys must be implemented.
- Panel alignment should be improved.
- It is hard to see the nested expressions.
- Hint must be improved.

From the comments, we think that the following comments are the advantage of LASP.

- Users can reduce effort generated by the hand.
- Users can reduce careless mistakes.
- It is easier to solve exercises than handwriting because users can try many tactics at short times.
- The time which we can take to learn by mistake is increased because users can reduce the efforts or time to substitute or inference.

4.4.2 Results of Solving Time

Table 6 shows the problems used in Experiment 1. Every problem is supposed that a thereom $P \rightarrow P$ has been proven.

Tables 7 and 8 show the results of Experiment 1. We withdraw the result of Exercise 3 because it involves a mistake. Table 7 shows the number of subjects who solve exercises correctly(S_c) and all subjects(S_{all}). Table 8 shows average solving time in which subjects correctly solve exercises.

Table 9 shows the problems used in Experiment 1. Every problem is supposed that a thereom $P \rightarrow P$ has been proven.

Tables 10 and 11 show the results of Experiment 2. Table 10 shows S_c/S_{all} and Table 11 shows average solving time in which subjects solve correctly exercises.

5 Discussion

First, we consider the usability which is evaluated in Experiment 1. We can see that users are unsatisfied with LASP's usability from the results of $Q1$, $Q2$ and $Q3$ in Experiment 1

Table 9: The problems in Experiment 2

	problem
$Q1$	$\neg A \rightarrow B \vdash \neg B \rightarrow A$
$Q2$	$\neg X \rightarrow (X \rightarrow \neg Y)$
$Q3$	$(\neg Y \rightarrow \neg X) \rightarrow ((\neg Y \rightarrow X) \rightarrow Y)$
$Q4$	$\forall x(A \rightarrow B(x)) \rightarrow (A \rightarrow \forall x B(x))$

Table 10: The number of subjects who solve correctly in Experiment 2

	handwriting	LASP
$Q1$	5/8	2/8
$Q2$	4/8	4/8
$Q3$	4/8	2/8
$Q4$	2/8	2/8

in Table 5. It is also shown by $Q6$ in Experiment 2. Therefore, LASP's user interface should be improved. In order to improve the interface, we have to consider the free comments. We think implementing shortcut keys and a feature which shows that users finish proving correctly are easy tasks.

Second, we consider efficiency of proving which is mainly evaluated in Experiment 2. From the results of $Q1$, $Q2$ and $Q3$ in Table 5, we can see that the substitution assistant feature and the inference assistant feature are contributed enough to the efficiency of proving. Users also feel that. In addition, twelve out of sixteen subjects comment that LASP can reduce effort by the hand in Experiment 2. Therefore, the goal of LASP's development seems to be enough achieved. The reason that the average point of $Q3$ is less than those of $Q1$ and $Q2$ seems to come from unsatisfaction of LASP's usability. Therefore, we should implement shortcut keys and other features.

From the results of Table 8 and 11, we can see that while proving by LASP is slower than by hand in Experiment 1, we obtained the opposite results in Experiment 2. The reason seems to be that while the time in which subjects can be familiarised with LASP is twenty minutes in Experiment 1, the time is an hour in Experiment 2. Therefore, we consider that if users are familiar with LASP then they can practice more effectively. Experiment 2 does not obtain enough samples; we should perform experiments in a large scale to prove the correctness of this hypothesis.

From the results of Tables 7 and 10, we can see that the number of subjects who can solve correctly exercises using LASP is the same number as ones who can solve correctly exercises by the hand. From the results, the effect of efficiency may not influence correctness of answering question.

Table 11: The average time of solving in Experiment 2

	handwriting	LASP
$Q1$	10m35s	5m57s
$Q2$	7m19s	6m04s
$Q3$	8m48s	5m54s
$Q4$	12m56s	5m19s

However, there is a mistake of applying inference rules and also two mistakes of substitution in handwritten answers. Using LASP, users can avoid careless mistakes because it shows an error message when users apply wrongly inference rules.

In conclusion, LASP is effective to learn formal proof.

6 Conclusion

This paper describes our learning assistant system for formal proof (LASP). LASP has features such as inputting exercise data, substitution assistant, inference rule assistant and so on. From the two experiments, we have obtained the results that LASP is effective though the user interface should be improved. Future work includes improvement of user interface and real application in the class.

Acknowledgments

REFERENCES

[1] David Abraham, Liz Crawford, and et al. A tool to practice formal proofs. *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications*, 2001.

[2] Christel Baier and Joost-Pieter Katoen. *Principles of model checking*. MIT Press, 2008.

[3] Dave Barker-Plummer, Jon Barwise, and John Etchemendy in collaboration with Albert Liu. *Tarski's World: Revised and Expanded*. Center for the Study of Language and Inf, 2007.

[4] Jon Barwise and John Etchemendy. *Turing's World 3.0: An Introduction to Computability Theory*. University of Chicago Press, Chicago, IL, USA, 1993.

[5] Jon Barwise and John Etchemendy. Computers, visualization, and the nature of reasoning. *The digital phoenix: How computers are changing philosophy*, pages 93–116, 1998.

[6] Yves Bertot and Pierre Castéran. *Interactive theorem proving and program development: Coq'Art: the calculus of inductive constructions*. Texts in Theoretical Computer Science. An EATCS Series. Springer-Verlag, 2004.

[7] Ana Bove, Peter Dybjer, and Ulf Norell. A brief overview of agda — a functional language with dependent types. In *Proceedings of the 22nd International Conference on Theorem Proving in Higher Order Logics*, TPHOLs '09, pages 73–78, Berlin, Heidelberg, 2009. Springer-Verlag.

[8] Bruno Buchberger, Adrian Crăciun, and et al. Theorema: Towards computer-aided mathematical theory exploration. *Journal of Applied Logic*, 4:470–504, 2006.

[9] Roy Dyckho. MacLogic: A Proof Assistant for First Order Logic on the Macintosh. *Computational Science Division, University of St. Andrews*, 1989.

[10] D. Gillet, Anh Vu Nguyen Ngoc, and Y. Rekik. Collaborative web-based experimentation in flexible engineering education. *Education, IEEE Transactions on*, 48(4):696–704, nov. 2005.

[11] Peter Henderson. Mathematical reasoning in software engineering education. *Communications of the ACM*, 46:45–50, 2003.

[12] Hannes Kaufmann and Dieter Schmalstieg. Mathematics and geometry education with collaborative augmented reality. *Journal of Computer and Graphics*, 27:339–345, 2007.

[13] Kiyoyuki Miyazawa. *LASP*, 2010. http://sdl.ist.osaka-u.ac.jp/~k-miyazw/.

[14] Charles Shubert, Ivica Ceraj, and Justin Riley. Bringing research tools into the classroom. *Journal of Computers in Mathematics and Science Teaching*, 28(4):405–421, 2009.

A Proposal and Evaluation of a High-Definition TeleVision Conference Service for Heterogeneous Clients with Various Signaling Control Protocols

Tsutomu Inaba†, Takashi Ogasawara†, Hideyuki Takahashi*, Naoki Nakamura‡, and Norio Shiratori*

†Business Marketing Division, NTT East Miyagi branch, Japan
‡School of Medicine, Tohoku University, Japan
* Research Institute of Electrical Communication, Tohoku University, Japan
{inaba, t.ogasawara}@miyagi.east.ntt.co.jp, nakamura@med.tohoku.ac.jp
hideyuki@riec.tohoku.ac.jp, norio@shiratori.riec.tohoku.ac.jp

Abstract - In order to reduce CO_2 emissions, "Green by ICT" are now attracting attention from various fields. HDTV conference system is one of this approaches, because movements of automobiles are inhibited by the use of HDTV conference systems. However, existing HDTV conference clients can not interconnect each other, if the clients are produced by different vendors, due to issues related to protocol. To solve these issues, this paper proposes a cloud-type interconnection service for heterogeneous HDTV conference clients, named *CISHDTV*, and evaluates it. The proposal is one of the results of "Kurihara Green Project", which is a subsidized project by the Ministry of Internal Affairs and Communications Japan. The CISHDTV is realized with Multi point Control Unit (MCU) and SIP conversion Gateway, and is provided to HDTV conference users beyond a public IP network.

Our demonstration experiments show a feasibility of the CISHDTV, and achieve a reduction of 80% CO_2 emissions in the case of a staff training in Kurihara-City, Japan, where several branch offices are placed widely. We are now discussing the CISHDTV in the multimedia sub working group of TTC in detail, towards an international standardization.

Keywords: HDTV Conference systems, SIP, H.323, MCU, Reduction of CO_2 emissions.

1 Introduction

"Green by ICT" is now attracting attention from various fields [1]. "Green by ICT" is an approach to inhibiting movements of automobiles/people, thus it is expected to reduce an effect on environments. "*Kurihara Green Project*" [3]–[8] in Japan is one of the projects promoted with "Green by ICT". The project especially forces on reducing CO_2 emissions in rural area instead of metropolitan area [9]. Rural cities such as Kurihara often have a distinctive structure, called "*Cluster-Type*" [10], in where several down towns are located separately. Hence, the people living in such a rural area have to depend on automobiles in their daily life. Therefore, it is important to limit the use of automobiles to reduce CO_2 emissions.

Television conference system is one of the approaches to reduce CO_2 emissions distinctly. Until now, several companies and/or communities have been introduced TV conference systems for the purpose of reducing the costs about business trip. However, existing TV conference systems produced by different vendors may not be interconnected with each other. The existing TV conference systems have several issues caused by a difference of communication protocols [2]. One is the issue caused by that the existing HDTV conference systems support multiple signaling protocols, such as H.323, SIP and SIP extensions. Another is caused by that the implementation of HDTV conference systems are different form one another, even though the systems are implemented based on same standardized protocol. And the other is caused by that each HDTV conference system supports different video codecs respectively. Because of these issues, a group which uses HDTV conference system has to prepare same products.

To solve these issues, this paper proposes "Cloud-type Interconnection Service for heterogeneous HDTV conference clients" (*CISHDTV*). The CISHDTV utilizes "Multipoint Control Unit" (*MCU*) and "Protocol Convergence GateWay" (*PCGW*). The service is provided for users beyond a public network on their demands. The service makes HDTV conference systems more convenient, and reduces CO_2 emissions significantly.

The rest of the paper is organized as follows. Section 2 explains the overview of the HDTV conference systems, and points out their interconnection problems. Section 3 proposes an interconnection service of heterogeneous HDTV conference clients beyond a public IP network. Section 4 evaluates the proposed service in terms of its feasibility and reduction of CO_2 emissions, and Section 5 concludes the paper.

2 Interconnection Issues in HDTV Conference System

This section describes about the technologies related to HDTV conference systems at first, and then points out the issues of their interconnections.

2.1 Technologies related to HDTV Conference Systems

Screen resolution of High-Vision (HDTV) is 1,920 x 1,080 pixels, thus it requires 786Mbps information transmission rate [13] to transmit images over networks. Therefore, HDTV conference system requires high video compression technique, because current internet access speed by optical fiber (FTTH) is 100Mbps or less. When a HDTV conference is held with more than three sites, MCU is required as a central apparatus for images and sounds.

HTDTV conference system is usually utilized on IP networks such as the Internet. Especially in a business use case,

the system is utilized on a Virtual Private Network (*VPN*) to keep the communication's security. Besides the above, the system also can be utilized on Next Generation Network (*NGN*) [11] or *HIKARI Denwa* [12]. NGN is a next generation communication infrastructure, which keeps the reliability and the stability that the traditional public telephone network provides. NGN also keeps the flexibility and the economic efficiency that IP networks provide. HIKARI Denwa is an IP based public telephone service. HDTV conference clients can connect each other over HIKARI Denwa with the 0AB J telephone number.

HDTV conference system generally employs either Session Initiation Protocol (*SIP*) or *H.323* as a signaling control protocol. The SIP can be used for creating, modifying and terminating two-party (unicast) or multiparty (multicast) sessions consisting of several media streams [14], [15] The existing SIP clients are generally implemented based on RFC3261 or RFC2543. Since the SIP defines only session control methods, protocols such as the Real-time Transport Protocol (*RTP*)[16] and Session Description Protocol (*SDP*)[17] are utilized to transmit not only video/voice packets but also information of media-ability. The HIKARI Denwa also uses the extended SIP, named *NGN-SIP*, which is enhanced from a viewpoint of securities. However, NGN-SIP does not allow some SIP methods which are generally used between HDTV clients and MCU, to prevent security incidents.

H.323 is a recommendation from the ITU Telecommunication Standardization Sector (*ITU-T*) that defines the protocol sets to provide audio-visual communication sessions on IP networks[18]. Since the recommendation of H.323 was firstly approved in 1996, a lot of H.323 clients have been appeared on the market one after another. H.323 has high affinity with the traditional public switched telephone network. However, because the protocol structure of H.323 is quite complicated, H.323 isn't suitable for internet technologies, and it can not scale. Therefore, the SIP is now the major signaling control protocol, instead of H.323.

2.2 Interconnection Issues of HDTV Conference Clients

2.2.1 The Connectivity between Different Signaling Control Protocols

An interworking technology between H.323 clients and SIP clients have been studied since the SIP was defined. For example, an equipment which helps interworking of these clients such as "Session Border Controller(*SBC*)" had been produced. As regards the requirements for the interworking technology, RFC4123 was defined as SIP-H.323 Interworking Requirements. However, media capability exchange-timing causes a serious issue in an interworking of two clients through the SBC. Each signaling control protocol defines original session initiation procedure, respectively. As shown in Table 1 and 2, H.323 and SIP have two session initiation procedures with different exchange-timings of media capability.

As shown in these tables, because H.323 clients and SIP clients support different exchange-timings of media capability, these clients can not connect each other in many cases.

Table 1: The capability exchange-timings in H.323.

	Capability Exchange-Timings
Fast Start	Client offers its media capability in H.225 / SetUp
Slow Start	Client exchanges its media capability with H.245 / Terminal Capability Set messages. Most H.323 HDTV clients support this method. (General initiation procedure of H.323.)

Table 2: The capability exchange-timings in SIP.

	Capability Exchange-Timings
Early Offer	Client transmits available codecs and bandwidth which is described in SDP of INVITE. Most of SIP HDTV clients support this method. (General initiation procedure of SIP.)
Delayed Offer	Client offers its media capability with 200OK, as a response of the Initial INVITE without SDP (no offer of capability).

It needs a combination of "Fast Start" and "Early Offer" or a combination of "Slow Start" and "Delayed Offer" to connect H.323 and SIP clients.

2.2.2 Difference between Implementations and Standardizations in the same Protocol

As described in Section 2.1, SIP and its extensions are defined in various RFCs. In addition, SIP relevant protocols, such as RTP and SDP, has been continually revised. These continually-updating standards make it more difficult to interconnect SIP clients, because fundamental standard of SIP client differs according to the timing of their development, even if they are developed by the same manufacturer.

Moreover, various implementations of many manufacturers prevent interconnection of clients. The SIP is a text based protocol, thus it can be implemented easily to conference systems. On the other hand, the text based protocol brings about ambiguities of SIP definition. For example, there is no restriction of characters in the SIP description. And the number of strings is not defined clearly. Due to these ambiguities, each manufacturer uniquely implements with their own interpretations. Therefore, the interconnection of SIP clients becomes more and more difficult. Similarly, an interconnection issue is also caused by the original SIP implementations in a communication network infrastructure. For example, several SIP clients are available with HIKARI Denwa. However, these clients can not be connected with each other, if these vendors are different from each other.

2.2.3 Difference of Implemented Video Codecs

An inconsistency in video codecs or communication rate often causes an interconnection issue of HDTV conference clients. The number of audio codec types is limited. Meanwhile,
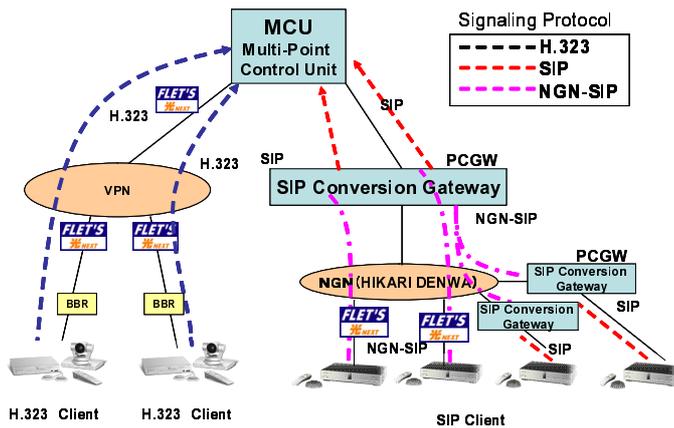
Figure 1: An overview of CISHDTV.

there are a lot of video codec types such as H.261 H.263 H.263+ H.264 and so on. Therefore, most of HDTV conference clients support a part of these video codecs based on their vendor's policy. As a result, HDTV conference clients can not be connected with each other, because the clients can not negotiate video codecs each other.

Similarly, a difference in communication rates supported by clients causes another issue of their interconnection, because each client specifies audio/video communication rate at its negotiation phase.

3 A Cloud-type Interconnection Service for Heterogeneous HDTV Conference Clients

To solve the issues as mentioned in Section 2, we propose a cloud-type interconnection service for heterogeneous HDTV conference clients, named *CISHDTV*, in which heterogeneous clients can be interconnected over public IP networks. Figure 1 shows an overview of CISHDTV. The CISHDTV can improve the users' convenience, in which the users can call a destination client on their demands with the 0AB J telephone number at a metered rate. The realization of temporal HDTV conference with any destination is greatly expected to stimulate a latent demand of users. As a result, HDTV conference gradually substitutes for traditional face to face meeting, and it greatly contributes to reduce CO2 emissions.

3.1 Interconnection between Different Signaling Control Protocols

We propose a MCU rental service beyond IP networks for the general public users on their demand. In the service, SIP and H.323 clients interconnect each other through the MCU. MCU generally adapts to various signaling control protocols, video codecs, transmission speeds, resolutions of picture, and so on. MCU can close not only the gap between SIP and H.323 but also the gap between the different abilities of various clients. Through an intervention of MCU, the issue of the capability-exchange-timing between SIP and H.323 as mentioned in Section 2.2.1 can be solved. In CISHDTV, the usage fee is charged based on the total length of time that the MCU service is used.

H.323 clients locate on a closed VPN to keep a conference session by H.323 clients secret, because H.323 clients can not be used on NGN and Hikari Denwa. NGN-SIP clients are directly connected to HIKARI Denwa network with its User-Network Interface (*UNI*), and call a destination client with the 0AB-J telephone number. SIP clients are connected to HIKARI Denwa through a Protocol Convergence GateWay (PCGW) which exchange NGN-SIP to SIP. CISDHTV service provider provides a socket connection interfaces to MCU with PCGW for SIP client users.

3.2 Translations for the Differences of SIP Implementation

We propose the method that PCGW bridges the difference of SIP implementation, to solve the issues as mentioned in Section 2.2.2. The PCGW can translate the RFC3261 based SIP into the NGN-SIP and vice versa. The PCGW can also bridge the differences of SIP implementation of clients. The PCGW supports both extension call and outside call. The existing SIP HDTV clients can connect to MCU over HIKARI Denwa with the PCGW. In CISHDTV, MCU is connected to HIKARI Denwa through the PCGW with 0AB-J telephone numbers. If a client calls to provider's PCGW with a phone number, the PCGW forwards the call to the suitable MCU, according to the configuration of call forwarding.

The existing HDTV conference clients generally use the SIP INFO Method [21] to transmit various information and control commands. However, the SIP INFO Method is not allowed to use in NGN/HIKARI Denwa (NGN-SIP) because of some security reasons. The PCGW encourages an interconnection of HDTV conference clients by its protocol translation function from SIP to NGN-SIP and vice versa.

3.3 Interconnection of Different Codecs

We propose the method that the MCU adjusts the difference of video codecs either communication rates that clients support, to solve the connectivity problem as mentioned in Section 2.2.3. MCU supports a lot of video codecs and communication rates, so as to connect various clients. Therefore, MCU adopts a video codec and/or a communication rate to the needs of the opposite client. For example, in the case where the caller A which supports video codec AA calls to callee B which supports BB through the MCU, the MCU negotiates A with AA and B with BB respectively.

4 Experimental Evaluations

In this section, we first verify the proposed CISHDTV in terms of clients' connectivity through demonstration experiments in Kurihara-city. Then we evaluate it for the reduction of CO2 emissions based on the environmental assessment methods.

4.1 Experimental Environments

Table 3 shows the list of equipments used in the experiments. We adopted a HDTV communication unit "Panasonic

Table 3: The list of used equipments

Category	Manufacturer	Equipment	Remarks
SIP client	Panasonic Communications	HD-COM (KX-VC500)	
NGN-SIP client	Panasonic Communications	HD-COM (KX-VC500)	Mode Switching
H.323 client	Sony Business Solution	PCS-XG80	
MCU	Cisco Systems andberg)	Codian4520	20Ports
SIP Conversion GW	NTT Software	Crossway	
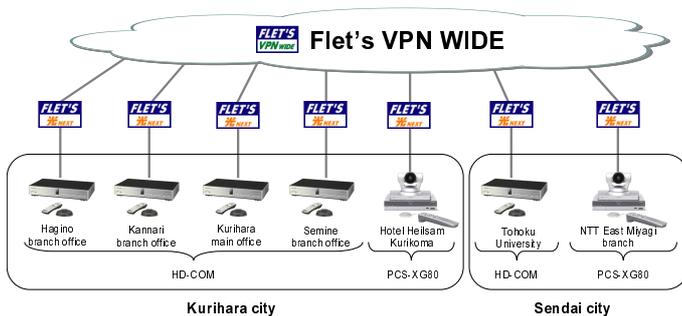
Figure 2: The deployment of HDTV clients.

Figure 3: Connectivity tests between SIP and NGN-SIP clients (STEP1).

1) Connectivity between SIP and NGN-SIP clients

2) Connectivity between SIP/NGN-SIP client and MCU

3) Connectivity between SIP/NGN-SIP client and H.323 client through MCU

In the experiments, we confirm the messages of signaling protocols during a conference session. We also measure and analyze the messages related with video and voice during a HDTV conference for over 30 minutes.

4.2.1 Connectivity between SIP and NGN-SIP Clients (STEP1)

At first, we try to connect HD-COM to each other with SIP and NGN-SIP mode. Figure 3 shows the pattern of connection. When HD-COM runs in SIP-Mode, HD-COM connects to HIKARI-Denwa through Crossway as a NGN-GW(PCGW).
[Case 1-1]
Caller HD-COM(SIP-Mode) calls to callee HD-COM(SIP-Mod) with IP address.
[Case 1-2]
Caller HD-COM(NGN-SIP) calls to callee HD-COM(NGN-SIP) with the 0AB-J telephone number.
[Case 1-3]
Caller HD-COM(NGN-SIP) calls to the opposite Crossway with the 0AB-J telephone number. In this case, the called Crossway with the 0AB-J telephone number calls the destination HD-COM(SIP) with the extended number associated with the 0AB-J number in advance.
[Case 1-4]
Caller HD-COM(SIP) calls to the opposite Crossway with the 0AB-J telephone number. In this case, the called Crossway with the 0AB-J telephone number calls the destination HD-COM(SIP) with the extended number associated with the 0AB-J number in advance. The caller HD-COM corresponds to an extension client of the Crossway.

The results from those four cases show that the interconnection of HD-COM with SIP/NGN-SIP modes does not cause any problems from the point of signaling control protocol

KX-VC500"(*HD-COM*) [22] which supports full HD quality in a conference. It can select from either SIP or NGN-SIP mode by switch. We also adopted a HDTV conference system "SONY PCS-XG80"(*XG80*) [23] which supports only H.323. Each client in this experiment is located as shown in Figure 2. The access line of each location is Flet's Next[24] to connect with NGN and HIKARI Denwa. H.323 clients are configured to connect to vpn service Flet's VPN Wide. The video codec of HD-COM is the H.264 High profile. Available resolutions of HD-COM are 1,920 x 1,080p, 1,280 x 720p and 704 x 480p. Besides HD-COM supports not only SIP but also NGN-SIP, so it can be used as a HIKARI Denwa client with the 0AB-J telephone number.

We also adopted the *Crossway* as a PCGW[19]. The Crossway is prepared for NGN-UNI (HIKARI Denwa) and NGN-SNI, thus we can use a existing HDTV conference clients on NGN through the Crossway.

Crossway realizes high connectivity, and it can connect with the policom, the tanbarg, the Sony and the HITACH high-tech clients. These four manufacturers occupy about 94% of share of SIP clients. The Crossway also can connect with the NGN-SIP clients such as the NEC and the Cisco. However the Crossway does not warrant the connectivity with HD-COMs.

4.2 Experimental Results

The evaluation experiments are executed in three steps below.

Figure 4: Connectivity test of SIP/NGN-SIP client with MCU (STEP2).

Figure 5: Image noise on the caller HD-COM

level and picture transfer protocol level. The results also prove that the interconnection between HD-COMs actually is very practical. It is clear that the Crossway can translate SIP of HD-COM into NGN-SIP and vice versa.

4.2.2 Connectivity tests of SIP/NGN-SIP client with MCU (STEP2)

In the STEP2, HD-COM in SIP/NGN-SIP mode connects to MCU. This experiment confirms that Crossway can translate the vendor original SIP implementations, and also confirms the connectivity of Crossway with various SIP clients in codec level.

[Case 2-1]
Caller HD-COM(NGN-SIP) calls to the MCU. Specifically, HD-COM(NGN-SIP) calls to the opposite Crossway with the 0AB-J telephone number, and the Crossway calls the MCU with the extended telephone number which is associated with the 0AB-J number in advance. In the case 2-1 and 2-2, we confirm the video pictures, which are fed back through the MCU, on the caller HD-COM.

The experimental results show that the interconnection of HD-COM works correctly from the point of Signaling Control Protocol level. However we also confirmed that image noise in Figure 5 was shown, sometime after a session is established.

[Case 2-2]
Caller HD-COM(SIP) calls to the MCU. Specifically, the Caller HD-COM(SIP) dials a 0AB-J telephone number to connect opposite Crossway, and the called Crossway calls the MCU using the extended telephone number associated with the 0AB-J number in advance.

These results show the interconnection between HD-COM and MCU works correctly from Signaling Control Protocol
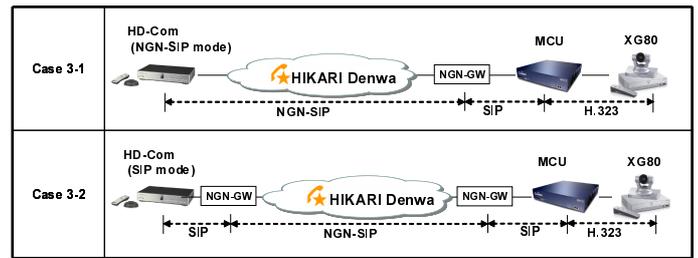
Figure 6: Connectivity tests of SIP/NGN-SIP client with H.323 client through MCU (STEP3)

level, and image noise does not confirmed at all.

4.2.3 Connectivity tests of SIP/NGN-SIP client with H.323 client through MCU(STEP3)

In STEP3, HD-COM connects to H.323 client XG80 through MCU. Here HD-COM and XG80 connect to the MCU respectively, according to an appointment information of meeting on the MCU.

[Case 3-1]
The caller HD-COM(NGN-SIP) dials a 0AB-J telephone number to connect to MCU beyond the HIKARI Denwa network. The XG80 dials an extended telephone number to connect to MCU beyond a VPN.

The experimental results show that the message connection between HD-COM and XG80 through MCU works correctly from the point of Signaling Control Protocol level. However we also confirmed image noise such as caused in case 2-1, sometime after a session is established.

[Case 3-2]
The caller HD-COM(SIP) dials a 0AB-J telephone number to connect to MCU beyond the HIKARI Denwa network. The XG80 connects to MCU with the same fashion of case 3-1.

The experimental results show that the message connection between HD-COM and XG80 through MCU works correctly from the point of Signaling Control Protocol level. In addition, image noise such as case 3-1 does not occur at all.

4.2.4 Discussions about the occurrence of image noise

In the test cases 2-1 and 2-2, image noise is appeared sometime after a conference session is established. Hence, we captured the messages transmitted between HD-COM and MCU with wireshark [20], and analyze these. As a result, it is clarified that the transmitted messages which HD-COM requests an opposite equipment to redraw its video image, does not arrive at MCU.

HD-COM requests the destination clients to redraw video pictures with RTP/AVP [25]. In fact, HD-COM transmits the RTCP messages to notify the destination clients that a video image has changed significantly. However, even though the Crossway received the messages from HD-COM, the Crossway could not forward the messages to MCU. Figure 7 shows the Initial INVITE message of SIP which is transmitted from HD-COM. In this figure, Media Description is described as "video 5264 RTP/AVP 109" in a part of Session Description

```
Frame 1: 900 bytes on wire (7200 bits), 900 bytes captured (7200 bits)
Ethernet II, Src: Panasoni_b1:37:60 (00:80:f0:b1:37:60), Dst: Sumitomo_e7:e6:2b (00:0b:a2:e7:e6:2b)
Internet Protocol, Src: 192.168.1.2 (192.168.1.2), Dst: 192.168.1.1 (192.168.1.1)
User Datagram Protocol, Src Port: sip (5060), Dst Port: sip (5060)
Session Initiation Protocol
  Request-Line: INVITE sip:03680186120ntt-east.ne.jp SIP/2.0
  Message Header
  Message Body
    Session Description Protocol
      Session Description Protocol Version (v): 0
      Owner/Creator, Session Id (o): - 766735355 1628862141 IN IP4 192.168.1.2
      Session Name (s): -
      Connection Information (c): IN IP4 192.168.1.2
      Time Description, active time (t): 0 0
      Media Description, name and address (m): audio 5132 RTP/AVP 9
      Bandwidth Information (b): AS:104
      Media Attribute (a): rtpmap:9 G722/8000
      Media Attribute (a): sendrecv
      Media Description, name and address (m): video 5264 RTP/AVP 109
      Bandwidth Information (b): AS:1005
      Media Attribute (a): rtpmap:109 H264/90000
      Media Attribute (a): fmtp:109 profile-level-id=42e01f;max-mbps=108000;max-fs=3600
      Media Attribute (a): sendrecv
```

Figure 7: Initial INVITE message from caller HD-COM(NGN-SIP mode).

Figure 8: A training workshop for officers of the Kurihara-city government with HDTV conference services.

Protocol(SDP). Although the Crossway supports only RTCP-Based Feedback (RTP/AVPF) [26], it does not support the RTP/AVP.

These results revealed that it is necessary for HD-COM to support RTCP-Based Feedback (RTP/AVPF) or for the Crossway to support RTP/AVP. We are now discussing which of two equipments is more suitable to adjust to the other.

4.2.5 The Effects of Reduction of CO2 Emissions

This subsection evaluates the effect of utilization of HDTV conference systems with CISHDTV from the view point of the reduction of CO2 emissions. As evaluation example in real world, we held "Training workshop of city worker". The workshop was held with thirty members of Kurihara City office, and twenty-one members from Kannari. In addition, two members from Hatsudai (Tokyo) also participated in the workshop as lecturers.

Table 4 shows an evaluation model and its assumptions. If all of the members gathers in a same conference room at Kurihara City office, twenty-one people from Kannari and two lecturers from Hatsudai have to visit Kurihara City office by automobile, subway and bullet train.

On the other hand, in the experiments, the members participated in the workshop with HDTV conference systems only needed to gather from their own offices. Figure 8 shows an overview of the workshop with HDTV conference systems. The XG80 (H.323 clients) in Hatsudai is connected to the MCU provided by CISHDTV beyond VPN. The HD-COMs (in NGN-SIP mode) in Kurihara and Kannari are connected to the MCU beyond NGN(HIKARI Denwa). Since the system in Hatsudai differs from the other systems, at least the lecturers have to go to Kurihara-city, if the CISHDTV is not provided.

The reduction effect of CO2 emissions are evaluated with *Kankyo Shiro* [28]. The Kankyo Shiro is an evaluation system to measure environmental impacts. It measures how ICT services over networks can reduce the amount of CO2 emissions quantitatively. The Kankyo Shiro is fully compliant with the standardized guideline of Life Cycle Assessment Society of Japan Forum [29] under Ministry of Economy, Trade and Industry. In addition it considers about every life cycle stage of ICT services, through its production, use and disposal.

This system automatically displays the amount of CO2 emis-

Table 4: Evaluation model and its assumptions

	Traditional Conference	HDTV Conference
Common Conditions	[Participants] Kurihara:30, Kannari:21, Hatsudai:2, [Meeting Time] 90 min	
Evaluation Model	All participants go to the meeting room in Kurihara city office.	Interconnect Kurihara, Kannari and Hatsudai.
Assumptions	[Traveling by Train] -Participants from Hatsudai travel back and forth between Hatsudai and Kurikoma-Kogen station. [Traveling by Automobile] -Participants from Kannari travel between Kannari-SogoSisyo to Kurihara city office by six cars. -Participants from Hatsudai travel between Kurikoma-Kogen station to Kurihara city office by a car. -Automobile mileage: 11.2Km/L	[ICT System] -Conference on 90 minutes -Scheduled 72 hours/year -The distance from Kurihara to Kannari: 12km -Flet's VPN Wide -Communication Traffic: 8GB/h

sions by the entries concerned about ICT System, Software, Movement of People, Movement of Goods, Computerization, Transportation Efficiency and Behavior of people. The Kankyo Shiro calculates the environmental load of the two cases, where ICT systems are introduced or not, and shows the effects of the case with ICT service compared with traditional method.

Figure 9 shows the reduction effect of CO2 emissions in this experiment. From the result of the experimentation, it is calculated the amount of reduction of CO2 emissions is about 29.46kg-CO2. This is roughly equal to the mount of seven day's CO2 emissions of average household [30]. As shown in Figure 9, the reduction rate achieves about 80%, this is because the amount of CO2 emissions caused by traffic movement exceeded that of by ICT systems.

Figure 10 shows the factors of CO2 emissions in this experiment. The movement of the lecturers by train has occupied about 58% of total CO2 emissions in the traditional confer-
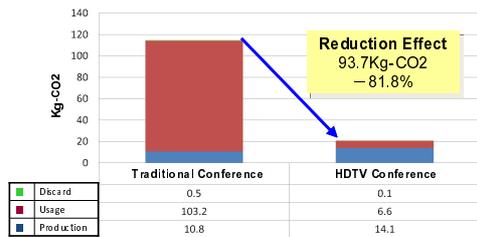
Figure 9: Reduction effect of CO2 emissions.
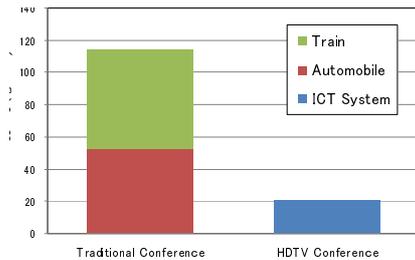
Figure 10: The factors of CO2 emissions.

ences. In reality, train service is independent of the lecturers' action. However, Kankyo-Shiro takes account of the CO2 emissions as a potential for reduction. Kankyo-Shiro calculates CO2 emissions with the basic data, which is defined as a unit amount in which a person moves 1 kilometer by train. This is because that Kankyo-Shiro is based on the premise, that a development and spread of IT technology makes people not to take trains and then reduces train services. Even if it does not take into account the movements by trains, we can see that the reduction of CO2 emissions achieves about 40%.

In the experiments, the officers of Kannari made a round trip to the Kurihara City office by five cars. The distance from Kannari to Kurihara is about 12 kilometers. The Kankyo Shiro calculated that the amount of CO2 produced by car movement is about 40kg-CO2. On the other hand, Kankyo Shiro calculates that the amount of CO2 will be produced by railroad is about 17Kg-CO2, if a railroad is built between Kurihara and Kannari. The reduction of CO2 emissions by using HDTV conference systems is more effective on clustered-type city such as Kurihara City, compared with metropolis with convenient public transportations.

5   Conclusions

This paper has proposed a cloud-type interconnection service for heterogeneous HDTV clients (CISHDTV) for the purpose of reduction of CO2 emissions. Demonstration experiments in Kurihara-city show a service feasibility of CISHDTV. However there are still problems left that messages instructing opposite-side client to refresh pictures could reach to the client when HD-COM client is in NGN-SIP mode. This causes the picture of the destination client severely distorted.

Now we are analyzing for the difference of implementation in detail, and are discussing the proposed service in Multimedia SWG meeting of TTC [31] toward its standardization.

In addition, demonstration evaluations in Kurihara-city verify the effectiveness of proposed CISHDTV, especially can achieve more than 80% reduction in CO2 emissions.

With the spread of the CISHDTV, a company or municipality which have already introduced the HDTV system come to exploit more opportunity to connect with other systems. As a result, it will come to have a great effect on the reduction of CO2 emissions.

In the near future, many organizations especially located around local city will use HDTV conference systems with CISHDTV.

REFERENCES

[1] S. Yoshizawa, "Approaches for reducing CO2 emissions by ICT Services," NTT technical journal Vol.4, pp.21–25, February 2011. (in Japanese)

[2] K. Singh and H. Schulzrinne, "Interworking between SIP/SDP and H.323," in Proceedings of the 1st IP-Telephony Workshop (IPtel 2000), (Berlin, Germany), April 2000.

[3] N. Shiratori, K. Hashimoto, D. Chakraborty, H. Takahashi, T. Suganuma, N. Nakamura and A. Takeda, "Kurihara Green ICT Project – Towards Symbiosis between Human ' s Life and Nature," Journal of Internet Technology(JIT), Vol.12, No.1, pp1–11,2011.

[4] T. Takahashi, M. Terabe, N. Nakamura, A. Takeda, T. Suganuma, K. Hashimoto, N. Shiratori, "Kurihara Green Project - Concept of ICT Systems toward Environmental load-reducing Town Development -," The 73rd National Convention of IPSJ, pp.3-17 - 3-18, Mar. 2011.

[5] T. Inaba T. Ogasawara Y. Seki H. Takahashi T. Suganuma, K. Hashimoto, N. Shiratori, "Kurihara Green Project - Life Support System to Reduce Carbon Dioxide Emissions -," The 73rd National Convention of IPSJ, pp.3-19 - 3-20, Mar. 2011.

[6] T. Inaba T. Ogasawara   H. Takahashi T. Suganuma, K. Hashimoto, N. Shiratori, "Kurihara Green Project - A Park & Ride Support System on Smart Phones -," The 73rd National Convention of IPSJ, pp.3-21 - 3-22, Mar. 2011.

[7] K. Matsumoto T. Sato H. Takahashi T. Suganuma, K. Hashimoto, N. Shiratori, "Kurihara Green Project - Green oriented Management Information Base (G-MIB) -," The 73rd National Convention of IPSJ, pp.3-23 - 3-24, Mar. 2011.

[8] M. Fujiwara T. Watanabe H. Takahashi M. Terabe T. Suganuma, K. Hashimoto, N. Shiratori, "Kurihara Green Project - Energy management system in wide earea

decentralization region -," The 73rd National Convention of IPSJ, pp.3-25 - 3-26, Mar. 2011.

[9] Hiroshi Esaki, "Eco-System Design Based on Internet Architecture Framework", EcoDesign 2009, Sapporo(Japan), December 2009.

[10] "Kurihara city Urban Transportation Master Planning," http://www.kuriharacity.jp/kuriharacity/contents/procedure/ doboku/toshikei/tko_master/index/1260276421/kotu_01.pdf/.

[11] "Next-Generation Services Joint-Development Forum", http://www.ngs-forum.jp/.

[12] "HIKARI Denwa Flet's public Home Page", http://flets.com/hikaridenwa/.

[13] S. Okubo, S. Kadono, Y. Kikuchi, and T. Suzuki, "H.264/AVC Textbook (in Japanese)," Impress Publisher, 2004.

[14] "SIP: Session Initiation Protocol," http://www.ietf.org/rfc/rfc2543.txt/

[15] "SIP: Session Initiation Protocol," http://www.ietf.org/rfc/rfc3261.txt

[16] "RTP: A Transport Protocol for Real-Time Applications," http://www.ietf.org/rfc/rfc3550.txt

[17] "SDP: Session Description Protocol," http://www.ietf.org/rfc/rfc2327.txt

[18] "H.323 Forum," http://www.h323forum.org/papers/

[19] "IP telephone Gateway Crossway," http://www.ntts.co.jp/products/ngn_appliance/

[20] "WIRESHARK the world's foremost network protocol analyzer," http://www.wireshark.org/.

[21] "RFC 2976 - The SIP INFO Method," http://www.faqs.org/rfcs/rfc2976.html

[22] "HDTV Video Conference System," http://panasonic.biz/com/visual/

[23] "HDTV conference system PCS-XG80," http://www.sony.jp/pcs/products/PCS-XG80/

[24] "FRET'S HIKARI NEXT," http://flets.com/

[25] " Extended RTP Profile for Real-time Transport Control Protocol (RTCP)-Based Feedback (RTP/AVPF)," http://www.ietf.org/rfc/rfc3551.txt

[26] " RTP Profile for Audio and Video Conferences with Minimal Control," http://www.rfc-editor.org/rfc/rfc4585.txt

[27] "State of the Environment (recycling society/life's diversity)," Ministry of the Environment　NIKKEI PRINTING INC, June 2010

[28] "LCA support tool for ICT services, Kankyo Shiro," http://eco.goo.ne.jp/business/keiei/ solution07/i/02.html/.

[29] The environment efficiency evaluation guideline for Information and Communication Technology (ICT) http://www.jemai.or.jp/japanese/eco- efficiency/pdf/guideline.pdf/

[30] Information Plaza of Electricity, The Federation of Electric Power Companies of Japan http://www.fepc.or.jp/present/jigyou/japan/index.html/

[31] "Telecommunication Technology Committee", http://www.ttc.or.jp/.

Adaptive Tabletop Dish Recommendation by the

Real-time Recognition of Dining Status

Tomoo Inoue[*]

[*]Graduate School of Library, Information and Media Studies, University of Tsukuba, Japan
inoue@slis.tsukuba.ac.jp

Abstract –This paper introduces a series of the systems named "Future Dining Table" which are the smart dining tables that recognize the user's dining activity and the food remains in real time, and presents its current version. The system recommends dishes to the user visually during dining by using above-mentioned information as the context of the user. The system is supposed to be useful where repeated dish orders take place while the number of serve staffs is not enough. A typical example field is a Japanese pub. Evaluations are given in the recognition accuracies of eating action and food remains, which confirmed the current system's practicality [1].

Keywords: Dining computing, Tabletop, Behavior recognition, Future Dining Table.

Figure 1: The system in use.

1 INTRODUCTION

Dining table, or the dining area, is a place where people get together and eat together. Regardless whether it is at home or is at public place, it is necessary for everybody because nobody can live without eating.

Information and communication technologies have become prevailing in many areas where we live. Dining table is one of the few places ICT would be deeply applied from now on.

We have been focusing on this research field as "Dining computing" since 2007. We have applied the analogy of Computer Supported Cooperative Work (CSCW) to categorize the field of dining computing. In CSCW, the setting is divided by time and space; "synchronous (real time)" or "asynchronous", and "face-to-face (local)" and "distributed (remote)." Because dining can be regarded as a kind of tasks, and dining together that involves multiple number of people can be regarded as a kind of collaboration, the setting of dining computing can be divided in the same way.

In this consideration, real-time face-to-face setting became a category of dining computing. Systems for this setting correspond to face-to-face electronic meeting systems. One of their basic functions is enhancing face-to-face meetings. A dining environment enhancing face-to-face dining can be a typical application system. A smart dining table that

understands the real-time context of the local users can be an application system of the category.

In real-time and distributed setting, systems for this setting correspond to remote desktop video conferencing systems. One of their basic functions is providing communication line when both users at different sites are in front of the cameras. A real-time tele-dining support system can be an application system in the dining computing.

Asynchronous and face-to-face setting calls for time shift communication or collaboration support. Many families hold members with different life patterns and they sometimes have difficulty in finding shared dining time. Application systems of this category may support time shift dining. It may use shared objects or something that sticks to the location because of face-to-face setting.

Asynchronous and remote setting calls for time shift remote collaboration support. This setting includes a family with some members live far distant with different time-zones. It has most relaxed condition for the users but supports least communication channels. Application systems may use interface agents that intermediate the remote users.

We call the dining environment that represents the field of dining computing "Future Dining Table," which includes all application systems of all categories. However not all the systems can be built at the same time. The adaptive tabletop dish recommendation systems that are presented in this paper are early application systems of dining computing for real-time face-to-face setting.

So far, we have developed a few versions of FDT. It has evolved gradually. This paper introduces those FDTs in Chapter 2 and Chapter 3. The system recommends dishes according to the dining status of the user (Figure 1). The system components are fairly simple. It has a USB camera
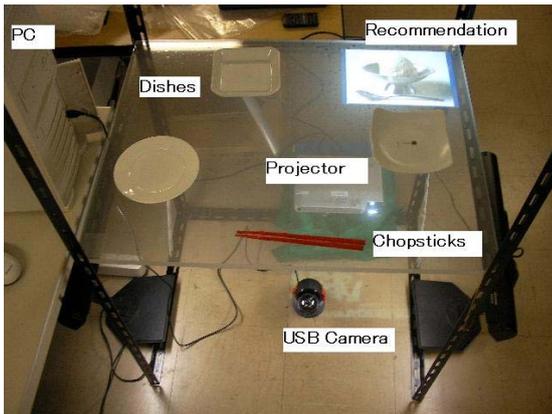
Figure 1: Appearance of the first FDT.

Figure 2: Visual markers on the bottom of the
dish and on the user's hand.

(a) Original image

(b) Extracted image

Figure 3: Hand recognition by image processing.

as the input device, a PC as the information processing unit, a projector as the output device, and a table for dining and for a screen of the information.

In Chapter 4, the system's recognition accuracy of eating action was examined and compared to the previous system with different recognition method. The system's recognition accuracy of food remains was also examined in Chapter 5. These evaluations were resulted in suggesting the practicality of the system.

2 FUTURE DINING TABLE

2.1 First Version

FDT is a tabletop system and recommends dishes according to the dining status of the user.

The appearance of the first version is shown in Figure 1 [1]. The table was a transparent 15mm thick acrylic board with 60cm depth and 75cm width, sealed by the transparent screen sheet for video projection so that the recommendation can be displayed on the table. Each dish on

the table had a visual marker on its bottom. The markers were recognized from the image taken by a USB camera under the table. The user also put the visual marker on his/her hand, which makes him/her feel unnatural (Figure 2). The recognition was not very robust with the roll of the wrist. The marker was sometimes occluded.

2.2 Second Version

To prevent these problems, the user's hand was recognized by image processing in the second version [2]. The hand was recognized by the background subtraction method. The image of the table was captured first as the background. To cope with the gradual change of the shooting condition, every frame was combined at the rate of 0.01 with the background. To cope with the change of the dish location and the change of the recommendation image, the background was updated. It was replaced by the foreground image when the foreground remained the same in 50 consecutive frames. Then the background was subtracted from the current frame, and the changed region was gained. After the opening to delete minor noise, the region with some area (more than 1000 pixels) was extracted by the labeling (Figure 3). This is the hand recognition process.

However, the same visual markers were still used on the bottom of the dishes with the transparent table, and the camera was under the table, which some users commented might be an issue.

Also, the food remains were not measured directly from actual food.

Figure 4: Proposed version.
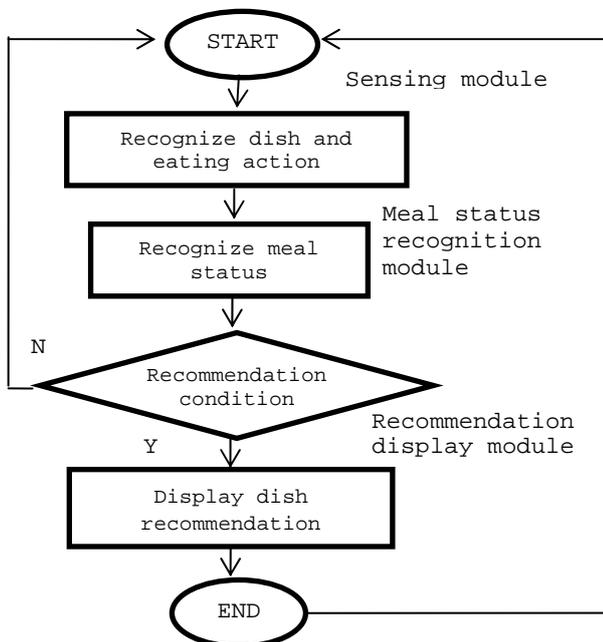
Figure 5: Software procedure.

2.3 Current Version

The current version has been modified in above-mentioned issues. A USB camera has been installed on the ceiling of the table and has recognized the dishes and the user's behavior. A white table has been used for the dining table with information display. A projector has also been installed on the ceiling of the table to project information on the table [3].

3 FDT SOFTWARE

The software of the current FDT is explained in this chapter. It has been implemented by Microsoft Visual C++ on Windows OS. Figure 5 shows the procedure, which is constructed by the sensing module, the meal status

Figure 6: Hand recognition from the ceiling.

recognition module, and the recommendation display module.

3.1 Sensing Module

The sensing module has been implemented with the Intel OpenCV. Four major processes are performed after obtaining the table image from the USB camera. They are hand recognition of the user, dish recognition, eating action recognition, and food remains recognition. When eating action is recognized, this triggers the meal status recognition module. When eating action is not recognized, the table image is replaced by its next frame.

The frame rate of the camera is 9 frames per second.

Hand recognition

The finger tip of the user becomes most distant from the user's body and becomes close to the dishes on the table when eating action takes place.

The hand is recognized by the background subtraction method. The image of the table is captured first as the background. To cope with the gradual change of the shooting condition, every frame was combined at the rate of 0.01 with the background. To cope with the change of the dish location and the change of the recommendation image, the background is updated. It is replaced by the foreground image when the foreground remains the same in 25 consecutive frames. Then the background is subtracted from the current frame, and the changed region is obtained. After the opening to delete minor noise, the region that is more than 1000 pixels is extracted by the labeling. Figure 6 shows the hand recognition where the hand moves 23.3 cm/sec.

Dish recognition

The dish recognition is performed by the colors of the dish rims. Base colors of the round dishes are white with 5 different colors on the rims. The colored regions are

Figure 7: Dish recognition by colors.
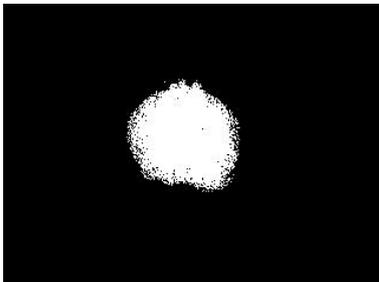
Figure 8: Food remains recognition.

extracted for each color and labeled to obtain the resulted region. Figure 7 shows the yellow dish recognition. The two points with the maximum and minimum x coordinates of the obtained region and the line between these two points can be gained. The line between the two points where the y coordinates are the maximum and minimum of the obtained region can be gained in the same way. The dish center can be gained as the intersection of the two lines.

Eating action recognition

Eating action recognition is performed by the distance between the dish center and the finger tip of the user. When the distance becomes less than the radius of the dish plus 10

Table 1: Example rules for the recommendation.

Status	Recommendation
Food remains is 25%	Closing food (Soba etc.)
Food remains is 5%	Desert (Ice cream etc.)
25% of 1st plate (Caesar salad etc.)	2nd plate (Gyoza etc.)
3 serial eatings of fries (Fried chicken etc.)	Same category plate (French fries etc.)
More than 6 drinkings (Beer) in the recent 10 eatings	Nibbles (Boiled soybeans etc.)

pixels, it is determined that the eating action occurs empirically. It is because the user holds chopsticks with his/her hand but the chopsticks are too thin to get the image. The user picks up the food with the distance. Usually the system has multiple dishes on the table. If the finger tip becomes less than the "eating action" distances for multiple dishes in this case, it is treated the food is taken from the nearest dish. However this does not really happen because with the determined distance it is not realistic for the hand to occupy such a position.

Food remains recognition

One of the method to recognize food remains is to estimate from the number of eating actions [1][4]. This method needs the information of the number of eating actions to finish the dish. Then the food remains is estimated as the ratio of the current number of eating actions to that of finishing the dish. Because the number of eating actions to finish the dish can vary depending on the various factors such as individual, food, health condition, the estimation is not very accurate. Another method is to measure the weight of the foods. The food remains is calculated as the ratio of the current weight to the initial weight [5][6]. This method can be accurate but is not very easy to implement because the weight of each dish must be measured. Weight scales may be embedded to the dining table.

Our research employs image processing method. This method estimates food remains from the 2D images of the dish. The result may not be very accurate due to the use of 2D images, but may be obtained fairly easily and is directly based on the actual food.

Figure 8 shows the result. Non-white pixels are counted from the white dish area. This is compared to the number of the initial pixels and the rate is regarded as food remains.

3.2 Meal Status Recognition Module

Meal status recognition module receives the food remains for each dish and the eating action history. Then the meal status is estimated from these. The food remains basically show the whole progress of meal, while the eating action history shows more precise chronological food consumption trend. Any rule to recommend a dish from the combination of the food remains and the eating action history can be made.

Figure 9: Dish arrangement in the evaluation
of eating action recognition.

Table 2: Eating action recognition of the current version

Subject	Precision	Recall	F-measure
A	.78	.88	.82
B	.96	.96	.96
C	1.00	.75	.86
Average	.90	.86	.88

Table 3: Eating action recognition of the first version

Subject	Precision	Recall	F-measure
D	.95	.79	.86
E	.82	.75	.78
F	.68	.54	.60
Average	.82	.69	.75

Example rules have been implemented in the system as shown in Table 1. "Rule 1: Food remains are 25%," and "Rule 2: Food remains are 5%" are derived from the survey result of dish recommendation timing. Recommended dish was most felt like placing the order at 25% and 5% food remains in the survey. Also recommendation was least disturbing at 25% and 5% food remains [7].

3.3 Recommendation Display Module

After the meal status is recognized and the recommended dish at the time is determined by the recommendation rules, the recommendation is displayed on the table. The recommendation display module shows the image of the recommended dish dynamically depending on the existing dish locations.

The existing dish locations are known from the dish recognition. The recommended dish location can be ruled in various ways and has been investigated. The locations near to other dishes and culturally right are felt easy to eat [8].

This module uses the result.

4 ACCURACY OF EATING ACTION RECOGNITION

4.1 Procedure

We have evaluated the accuracy of eating action recognition of the system. The same procedure was applied for the comparative evaluation to the first version which used the visual marker for the hand recognition and the current version which uses image processing for the hand recognition.

The subjects were 3 right-handed male university students. All were the first time users of the system. They were instructed to eat normally as in everyday. 3 dishes were set on the table as shown in Figure 9. A snack of small pieces was chosen as the food on the dishes because we could control the number of eating actions by this. 8 pieces of the snack were set on each dish. This meant the subject conducted 24 actions. The behavior in the session was videotaped. The record of recognition by the FDT was

compared with the videotaped behavior, which provided the correct answer.

4.2 Result

The results are shown in Table 2 and Table 3. Precision, recall and F-measure were used as the measures. These are originally from information retrieval, and have become generally used as the measures of such evaluation. Precision is defined as true positive / true positive + false positive. When an eating action from a false dish is recorded by the system, precision decreases. Recall is defined as true positive / true positive + false negative. When the system overlooks an eating action and does not record the action, recall decreases. F-measure is defined as the harmonic mean of precision and recall; $F = 2 / (1/Precision + 1/Recall)$. It is used to represent the performance of both precision and recall in a single measure.

Because this type of evaluation is unique, we cannot discuss the results in comparison with those of other research. The precision was .90 and the recall was .86 on the average, resulting .88 of F-measure in the current version. Whereas the precision was .82 and the recall was .69 on the average, resulting .75 of F-measure in the first version. The change of hand recognition method succeeded in improving the recognition rate.

5 ACCURACY OF FOOD REMAINS RECOGNITION

5.1 Procedure

We have also evaluated the accuracy of food remains recognition of the system.

The subjects were 3 right-handed male university students. They used FDT for dining. A single dish was served at one time. 2 different dishes were served as shown in Figure 10.

Table 4: Food remains recognition for boiled soybeans (%)

Subject	75%	50%	25%	5%
A	60.0	40.0	10.0	0.0
B	60.0	40.0	20.0	0.0
C	75.0	35.0	10.0	0.0
Average	65.0	38.3	13.3	0.0

Table 5: Food remains recognition for fried rice (%)

Subject	75%	50%	25%	5%
A	53.3	46.7	30.7	5.3
B	62.7	40.0	20.0	1.3
C	73.3	60.0	37.3	9.3
Average	63.1	48.9	29.3	5.3

Figure 10: Dishes used for food remains evaluation. (Top) Boiled soybeans. (Bottom) Fried rice.

One was "boiled soybeans," which each piece was not very small and the number of pieces was countable. The quantity was measured in pieces with this dish. We set 20 pieces initially and decreased by 5% for every eating.

The other was "fried rice," which had different appearance from the former thus was expected to produce a different result. The quantity was measured by weight with this dish. A cooking scale with the minimum scale was 5g was used for weighting.

Both dishes were common in the supposed environment of the system. Each subject ate up these 2 dishes.

5.2 Result

The results are shown in Table 4 and Table 5. The remains recognition for fried rice was more accurate than that for boiled soybeans. The result of 5% in Table 4 represents the recognition character clearly. The remains are recognized by a 2D image from the top. When the remains are recognized as 5%, the pixels of food are 5% of the initial pixels. This is only achieved when no overlapping of the food is found even if extraction of the pixels is accurate, and is often not realistic. In the experiment, only one piece of soybeans pod was 5% but the pixels for the one soybeans pod was clearly more than 5% of the initial pixels because the soybeans were piled up with overlapping on the dish. This is why the recognition of 5% always came after the actual 5%.

The reason why the difference in 75% was the biggest in Table 5 can be explained with the same recognition character. Fried rice is also piled up on the dish. When the subject starts eating, he does not usually eat fried rice by clearly finishing portion by portion like eating a pizza. Instead, he often eats some upper part without finishing to the bottom. Because of the 2D image recognition, height decrease is not recognized, and the remains are the same in this case as a result. Thus actual quantity becomes less than the recognized quantity by 2D image.

However, the differences between the recognized remains and the actual remains were within 5% when measured by the weight. Those two values were in good co-relation.

6 CONCLUSION

First, the short introduction of the concept of Dining computing was given.

Second, a smart dining table system "Future Dining Table (FDT)" that has been developed iteratively for versions as one of the realization of Dining computing was explained. The current system recognizes the user's dining activity and the food remains in real time, and recommends dishes to the user visually during dining along its context.

Then the evaluations of the recognition accuracies of eating action and food remains were explained, which indicated the system's practicality.

According to the categories of the dining computing and other demands such as communication support, FDT will be further extended in the future.

REFERENCES

[1] Y. Seto, Y. Noguchi, M. Tosaka, and T. Inoue, Development of Another dish recommender based on dining activity recognition, IEICE Technical Report, Vol.107, No.554, pp.55-60, 2008. (in Japanese)

[2] Y. Seto, Y. Matsusaka, and T. Inoue, Realtime dining activity recognition in Another Dish Recommender, IPSJ SIG Technical Reports, Vol.2009, No.3, pp.1-6, 2009.

[3] Y. Kourai, Y. Otsuka, and T. Inoue, Adaptive tabletop dish recommendation system by the recognition of realtime dining activity, IPSJ SIG Technical Reports, Vol.2010-GN-77, No.18, pp.1-8, 2010. (in Japanese)

[4] Y. Dong, A. Hoover, and E. Muth, A device for detecting and counting bites of food taken by a person during eating, IEEE International Conference on Bioinformatics and Biomedicine 2009, pp.265-268, 2009.

[5] T. Kawashima, Y. Tanisugi, and Y. Mitsudo, Dining monitoring system using sensing tray and ID-ware, IEICE Technical Report, Vol.106, No.285, pp.61-66, 2006. (in Japanese)

[6] K. Chang, S. Liu, H. Chu, J. Hsu, C. Chen, T. Lin, C. Chen, and P. Huang, The diet-aware dining table: Observing dietary behaviors over a tabletop surface, LNCS, Vol.3968, pp.366–382, 2006.

[7] Y. Seto, and T. Inoue, Study of recommendation of dishes for a dish recommender system, IEICE Human Communication Group Symposium, HCG2009-C5-1, 2009. (in Japanese)

[8] Y. Otsuka, and T. Inoue, Comparative study of attractive dish arrangements for a tabletop dish recommendation system, 2011 IEEE International Conference on Systems, Man and Cybernetics, 2011. (to appear)

Extending the Multi-user Tabletop Interface by Direction-aware Physical Objects for Collaborative Work in Planning Theater Production

Yosuke Horiuchi[†], Tomoo Inoue[‡], and Ken-ichi Okada[†*]

[†]Graduate School of Science and Technology, Keio University, Japan
[‡]Graduate School of Library, Information and Media Studies, University of Tsukuba, Japan
* Japan Science and Technology Agency
{yosuke, okada}@mos.ics.keio.ac.jp
inoue@slis.tsukuba.ac.jp

Abstract - Theater is a composite art that consists of various artistic expressions. Because theater involves various elements, such as the stage design and the sound effect, theater producers are required to create a picture of a theatrical stage as a whole in their mind. But it is hard to imagine the comprehensive theatrical stage without actual reproduction of it. In this paper, for the multiple users' collaborative work to plan a theatrical production, we propose a system that allows to actually reproduce theatrical productions with using a tabletop interface. The proposed system allows users to reproduce productions required for theater—the position of an actor/actress, the effect of sound and lighting—on a miniature theatrical stage which is reconstructed on the tabletop surface. The system allows users to use these system functions with two simple operations; touching the surface and moving an actual doll on the surface. In order to actualize the intuitive interaction with using a physical doll, we have developed the method of physical object recognition which enables the system to detect the doll's type, position and direction when it is placed on the surface. By reproducing common theatrical production elements with simple operations, users are able to call up an image of the theatrical stage as a whole. Through conducting user studies, we verified the effectiveness of the proposed system in planning a theatrical production.

Keywords: tabletop interface, tangible interface, multi-user, theater

1 Introduction

Theater is a composite art which is collaboratively composed with various artistic expressions. Because contributing the development of communication skills and a wealth of the expressiveness to workers, the creative activities in theater have been used as an educational methods [1]. But it is difficult to form a mental picture of a comprehensive theatrical stage which includes various theatrical productions, such as the effect of lighting and the stage design. From [2], to plan lighting production is hardly accomplished without having adequate study about them. In addition, for production members who are in different departments to plan each production, it is also hard to share their idea of theatrical productions to collaboratively create a theater. Although some applications have supported theatrical production planning, most of them have used a PC screen to express the theatrical stage, causing the difficulty imagining the three-dimensional

theatrical stage. In addition, most systems are made for individual work, therefore cooperative production planning with multiple producers is not considered.

In this paper, we propose a tabletop system for multiple users' collaborative work to plan a theatrical production that assists users to perceive the theatrical stage as a whole. In this system, the tabletop surface is used as a miniature theatrical stage on which some actors/actresses act. On the miniature stage reconstructed on the table, users can actually reproduce requisite theatrical productions; users reproduce the effect of lighting and sound by touching an operation window projected on the tabletop surface, and reproduce the position of the actor/actress by moving physical dolls as the tangible interface on the surface. In order to actualize the intuitive tangible operation with physical dolls on the DiamondTouch table [3], we have developed the method of physical objects recognition with which the system recognizes the doll's type, position and directions on the surface. By means of these functions, the system helps users get an image of the theatrical stage as a whole.

2 Creative activity in theater

Creative activity in theater is composite art that is coordinately created from various artistic expressions, such as sound, spatial design. This activity is proposed as the instruction method that promotes the development of the expression of feeling and the communication skills in school setting.

2.1 Process of theatrical production planning

Theatrical production involves various works. The following is an example of workflow in a producing activity in theater.

1. Decide on a script
2. Decide on overall theme
3. Decide on producers and performers
4. Performance rehearsal and planning theatrical production (such as the effect of lighting and sound, the costume and the stage design.)
5. On-stage rehearsal
6. Performance

In this research, we especially focus on theatrical production planning, which includes coordinated discussion and determination of productions. Generally, productions are made

by organizing works of each staff in charge of each production element, such as lighting staff and sound staff [4]. In this task, each production element must be designed based on a common vision shared among all staffs. Lack of shared vision makes the entire theatrical space mood spoiled.

2.2 Problems in theatrical production planning

The work of theatrical production planning includes some difficulties as follows.

2.2.1 The difficulty of visualizing the entire theatrical space

When initially planning theatrical production, each producer has to imagine the theatrical stage situation without having use of an actual stage. In this task, producers have difficulty imagining the entire stage only with utilizing each producer's capacity for imagination. As for lighting production planning, for instance, it needs a lot of effort to acquire enough skill to plan it. Therefore producers who have not enough skill are unsure what to do in planning [2].

2.2.2 The difficulty of sharing ideas of each production with other producers

In theatrical production planning, some producers in charge of different element cooperatively do the work while sharing their ideas with each other [5]. But even visualizing the entire theatrical space in each producer's mind is difficult, therefore sharing their pictures with others is obviously difficult.

2.2.3 The difficulty of creating a cue sheet

A cue sheet is a chart that shows the timing and effects of each theatrical production. In actual performance, staffs reproduce productions by using this sheet as a reference. Creation of this is cumbersome and difficult even if workers have enough experience in theater.

2.3 Related work

Some researches support creative activity in theater by promoting imagining the theatrical stage by displaying a virtual space which are created by computer graphics on a PC screen. Wu et al. [6] have designed a system that contributes a sensation of acting in theatrical stage by reflecting user's real-space action to virtual theatrical space. Bowen Virtual Theater [7] have assisted users to imagine a theatrical stage by depicting a stage sets and lighting production on a PC screen. There are also many commercial systems to design theatrical production. WYSIWYG [8] is used to plan, design and program lighting production. Matrix3 [9] supports the planning of sound production. These systems are intended for individual work, and have not considered the collaborative work which involves multiple users.

Avatar [10] is a system that supports multiple users' theatrical production planning by using virtual reality which represents certain theater scenes. However, being designed for work for multiple users away from each other, the system has not been intended for face-to-face collaborative work, which facilitates users' sharing of the elements, such as work objects and information.

Because most of researches have been intended for individual work as described above, we have developed a system [11] for multiple users' work to plan a theatrical production. The system helps users to imagine theatrical space with tabletop interface. On the tabletop surface which is regarded as the theatrical stage, the system allows users to reproduce theatrical productions. With this system, however, the reproducibility is too scarce to imagine the actual theatrical stage. In addition, because each user's working area is locked to reduce interfering in other users' work, users have difficulty in free operation.

3 System design

3.1 System concept

In this research, we aim to design a system that supports theatrical production planning for multiple users by promoting their work to perceive the entire theatrical stage. To actualize this, the system contributes a work environment that allows to use miniature model of theater. As pointed out by C.Dompierre [10], this method has been used for theatrical production planning until now.

In addition, the system aims to support for multiple users' collaborative work. When planning a theatrical production, some staffs who are in charge of different production cooperatively confer and make a decision [5]. Therefore, the system has to be designed to support multiple users' collaborative work.

For these, the tabletop interface appropriate for multiple users' collaborative work is used as work environment. By putting some miniature model of theater, such as dolls and miniature stage sets, the system reconstructs a miniature theatrical stage on the tabletop surface.

In this research, for multiple users' work to plan a theatrical production, we design a system based on the following concept to solve the problems in theatrical production planning.

3.1.1 To help users visualize a comprehensive theatrical stage

To actualize checking and imagining certain theatrical stage situation, the system allows users to actually reproduce their ideas of theatrical productions on the miniature stage. In addition, to check the positions and actions of the actor/actress, the system allows to use a physical doll as the tangible interface which is a surrogate for the actor/actress. By allowing reproduction of theatrical productions in three-dimensionally-formed miniature stage, the system assists users to visualize the theatrical stage as a whole.

3.1.2 To help users share a picture of the theatrical stage with other users

When planning a theatrical production, as previously mentioned, staffs are required to share their production ideas and

a common vision of production. With this tabletop system which contributes face-to-face work environment, users can directly exchange their opinions with others and share their production ideas by actually reproducing them on the shared miniature stage. Therefore, multiple users become easier to share a picture of a theatrical stage with others.

3.1.3 To facilitate a creation of a cue sheet

The system aims to cut back user's workload in creating a cue sheet, which is so cumbersome and takes so much time to create as described above. To actualize this, the system stores the reproduced productions during planning phase, and automatically generates a cue sheet from stored data.

3.2 A system for reproducing theatrical productions

On the basis of above concept, we propose a tabletop system that supports multiple users working collaboratively to plan a theatrical production. With this system, multiple users reproduce a theatrical production on a miniature theatrical stage reconstructed on the tabletop surface. The system also helps users create a cue sheet, which is so cumbersome that producers have difficulty creating it, by automatically generating it.

In order to help users visualize the entire theatrical stage, the system allows three of the production elements required for theater—the position of the actor/actress, lighting effects and sound effects—to be reproduced in a three-dimensionally-formed miniature theatrical space. By actually visualizing the ideas of theatrical productions on the miniature stage, users can create a clear picture of the theatrical stage in their mind and easily share their ideas of theatrical productions.

To reduce users' workload of creating a cue sheet, the system automatically generates it. With this system, for a certain scene, users are able to store the productions reproduced on the miniature stage. From stored data, the system generates a cue sheet by organizing them in an easy-to-understand format.

The system allows users to use these functions by only two simple operations; touching the tabletop surface and moving dolls on the surface. In addition, the system contributes a personalized feedback to each user even when multiple users are simultaneously performing individual operation on the same table. By using the system function with these intuitive operation, users can get centered on theatrical production planning.

4 Implementation

4.1 System configuration

As shown in Figure 1, the proposed system contributes the work environment for three users' collaborative work to plan a theatrical production. The system allows users to reproduce theatrical productions on the miniature theatrical stage reconstructed on the tabletop surface. In this research, we have used the DiamondTouch table [3] as the tabletop interface, which detects not only the location of the touch but who touches the

table surface while multiple users are simultaneously working on it. When users touch the table, the DiamondTouch gets such information by detecting electrical current flowed from users' body sitting on seats which are connected with the DiamondTouch. Therefore, three users touch the surface and reproduce theatrical productions while sitting on the seats. This system also allows to operate with physical dolls placed on the miniature stage by bringing in the physical object recognition method we have developed. To actualize this, the metallic plates to carry the electrical current flowed from users are stuck to dolls.
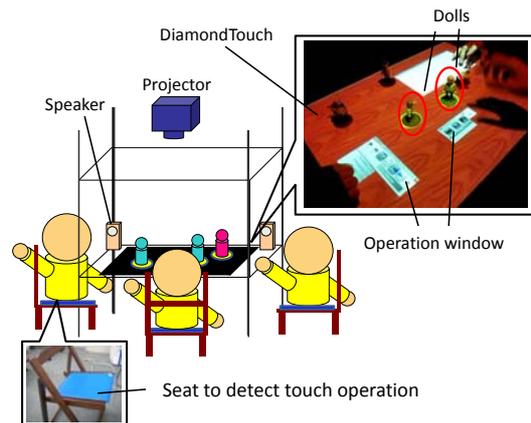
Figure 1: System architecture

4.2 Physical object recognition

In this system, users can use physical dolls to plan the position of actors/actresses. To operate with physical dolls on the table, we have developed the physical object recognition method for sensing electrical capacitance. In this method, by carrying an electric current through metallic plates attached to the bottom of each object, the system detects object operation on the table.

Under the DiamondTouch table surface, lattice-shaped electrodes are embedded at 5 mm intervals. When users sitting on the seats connected with the table touch the surface, the electrodes embedded under the touch position detect the electrical current flowed from each user's body. By using the electrical values detected by the electrodes, the system detects the object's type, position, and direction.

In this method, three metallic plates—the Base Plate (BP), the Identification Plate (IP), and the Direction Plate (DP)—are attached to the bottom of each doll in a double-layered arrangement, as shown in Figure 2. Then, the bottom of each doll is coated with conductive spray. When users place the dolls equipped with these metallic plates on the DiamondTouch table, the system detects the electrical current flowed from only the DP and the IP.

When detecting an input, the system firstly judges whether the input is caused by finger touch or doll placing. When judging, the system relies on following criteria which the doll placing input meets;
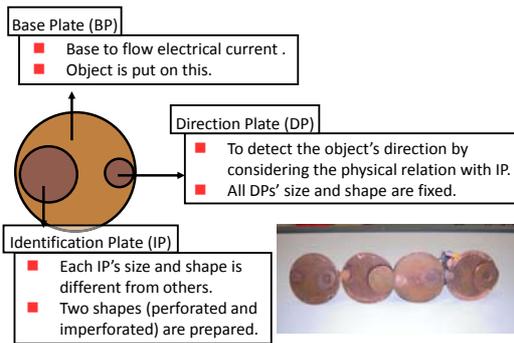
Figure 2: Double-layered structure with three plates

- The number of segments is one or two.

  Because the positional relation of the IP and the DP, the number of segments consists of reacted electrodes becomes one or two, as shown in Figure 3,5. So if not, the input is judged as the finger touch.

- The distance between two segments is not so large.

  The position of the IP and the DP are fixed with maintaining a certain distance. Therefore, if the distance between detected two segments becomes larger than the baseline, the input is regarded as the finger touch.

- The number of the reacted electrodes included in the detected segment do not exceed specified value.

  The each plate size is specified beforehand. Therefore, if the detected segment includes supernumerary reacted electrodes, the input is judged as the finger touch.

- The value of the noises between two segments exceeds specified value.

  Because a metallic plate is thin, the system detects slight electrical current flowed from the BP when the doll is placed on the tabletop surface. If the noises between two segments is not exceed specified value, the system judges the touch is caused by a finger, as shown in Figure 3.
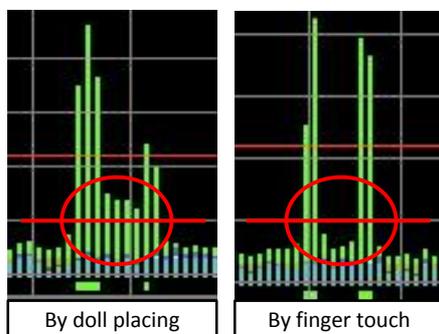
Figure 3: The use of the noise generated between two segments: if the input is caused by doll placement, the system detects slight electrical current flowed from the BP.

After the input is judged that it is caused by doll placing through conducting above checkup, the system detects doll's information in detail.

The system detects the type of the doll depending on their IP shape and IP size. In addition, the system detects not only their position but also their eight directions depending on the positional relation of IP and DP.

To prevent the false recognition between each object caused by the positional relation of IP and electrodes arranged at intervals of 5 mm, the each IP size attached to each doll must differ in increments of 1 cm. In this system, two types of IP shape was prepared, perforated and imperforated. Figure 4 shows three cases of the expected value of the reacted electrodes when the same size IP is placed on the DiamondTouch table. As shown in the figure, each shape of IP marks differences in detected electrical value, then the system can distinguish them.
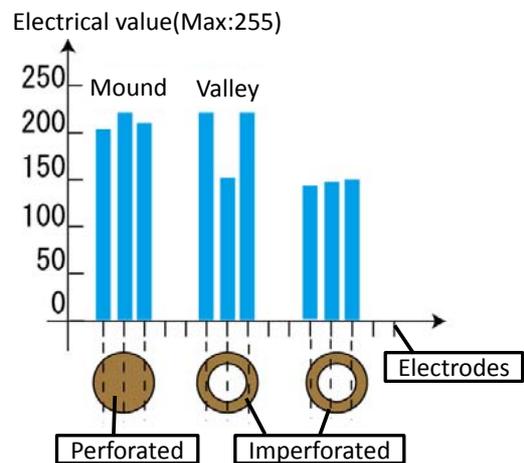
Figure 4: Three cases of the expected electrical value when the same IP is placed on the surface.

In addition, by utilizing the shape of reacted electrodes which shifts depending on the positional relation of IP and DP, the system detects eight directions for each doll, as shown in Figure 5.

Although there are some systems that employ tangible manipulation on a tabletop surface, for multiple users' work to plan a theatrical production, it is important to allow users to use physical objects on the DiamondTouch table, which recognizes who touches where on the surface. By actualizing this, the proposed system contributes the two system features which both are significant for the work: to give each user individual information which is appropriated for each user's role and to allow users to use the dolls on the table.

Being available in tabletop interface equipped with capacitance sensor, this method is utilizable for a variety of tasks, not only theatrical production planning. In this work, as shown in Figure 2, we prepared four dolls whose IPs are classified into two sizes and two shapes.
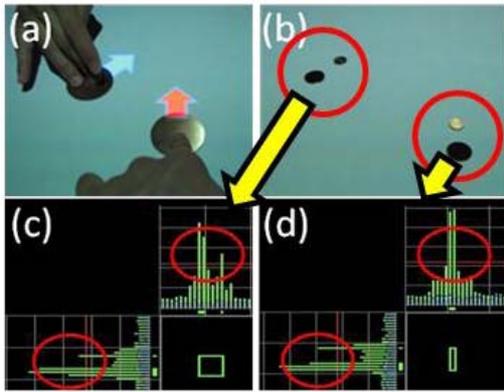
Figure 5: Direction recognition of (c) upper right and (d) up. The arrangements of reacted electrodes is differ depending on the positional relation of IP and DP.

4.3 Reproduction of theatrical production

To help users visualize the theatrical stage as a whole for theatrical production planning, the system allows three users—taking role of the stage director, person in charge of sound effect, and person in charge of lighting effect—to reproduce theatrical productions on the miniature stage. The system allows to reproduce three theatrical productions which are required for theater: the sound effect, the lighting effect, and the position of actors/actresses. To reproduce these productions, each user uses each individual operation window projected on the surface (as shown in Figure 6), which is dedicated for conducting each role. When reproducing and planning theatrical productions, each user receives individual feedback from the system by utilizing the feature of the DiamondTouch table that it detects who touches the table. When an user firstly touches the surface, for example, the operation window for the role of the user automatically appears at the touched position in an easy-to-understand direction, as shown in Figure 6. The system function of giving individual feedback reduces interference between each user, and then makes multiple users' collaborative work smooth.

Function of each operation window is described as follows.

- Stage director's window

  With this window, the user taking role of the stage director shows the script window and generates cue sheets. The script window includes many scenes consisting of dialogues and stage directions, which indicates a movement of an actor/actress and so on, for which users plan theatrical productions. Users can select the scene to store theatrical productions by touching it directly. While a production is stored to a certain scene, an icon to check detail of stored production is shown above the scene in the script window. When the user touches the icon of generating cue sheets, from the stored theatrical productions, the system creates cue sheets of sound and lighting in easy-to-understand format, as shown in Figure 7.

- Sound effect operation window

With this window, the user in charge of sound effect can conduct the control of sound production: playback, volume level, and fade-in and fade-out. The sound selected by the user is come from the speaker. By touching the OK icon, the user can store reproducing sound production to a selected scene.

- Lighting effect operation window

  The user in charge of lighting effect can do the control of global lighting and spotlight with this window. The available actions for both lighting productions are turning on/off (in case of spotlight, the user is also required to select the doll to focus on), setting the brightness and color. These actions for lighting production are reflected in the image projected on the tabletop surface by the projector set over the tabletop system. By touching the OK icon, the user can store reproduced lighting production to selected scene, as with the sound effects.
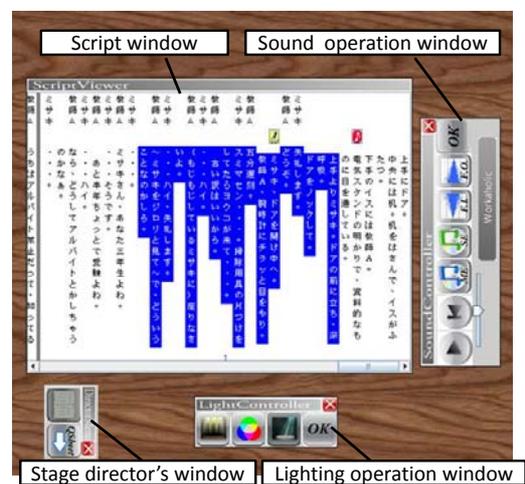
Figure 6: Operation window projected on the surface.

The operation of moving dolls is allowed for all users. Because the system can detect the location of each doll, the spotlight focusing on each doll automatically follows it even when it moves on the table.

4.4 Generation of a cue sheet

To help users who have *the difficulty of creating a cue sheet*, the proposed system automatically creates it by organizing the stored production information. The sound cue sheet includes each sound production information: the title, volume, start and end scenes, and effect such as fade-in and fade-out. The lighting cue sheet includes: the lighting type (global lighting or spotlight), brightness, color, and start and end scenes. As shown in Figure 7, in the sound cue sheet, each sound production information is showed in graph form as well as in numerical form. In order to enable users to easily recognize the sound production which each graph explains, each graph is expressed by a color identical with a background color of each production number list shown at the top of the sound cue

sheet. The graph explains each sound production information: the volume is expressed by the graph's height, fade-in and fade-out are expressed by the graph's slope, and the start and end scenes are indicated by the graph's range coordinated with the scene cell under the graph. By visualizing sound production information as graphs, the cue sheet becomes easy-to-understand format.

6	5	4	3	2	1	Qnumber
ME"workaholic"	SE"rain"	SE"wind"	ME"inTheLakeSide"	SE"thunder"	ME"cheery"	Sound
FI	FI	CI	FI	CI	FI	Start
40	30	90	80	50	60	MaxVolume
FO	FO	FO	CO	CO	FO	Finish
3 2 2 2 2 1 1 1 1 1 1						Page
（シーン：縦書きテキスト）						Scene
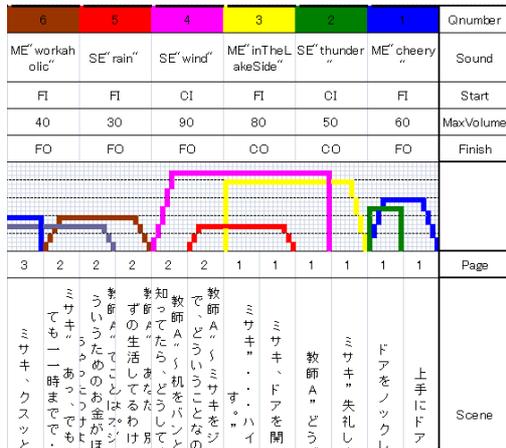
Figure 7: A sound cue sheet generated automatically.

5 User study

We conducted two evaluation experiments to clarify the effectiveness of the proposed system for multiple users' theatrical production planning.

5.1 Verification experiment for effectiveness in visualizing the entire theatrical stage

The goal of this experiment is to investigate the proposed system's contribution for helping users who have *the difficulty of visualizing the entire theatrical space*.

In this experiment, the participants actually planned theatrical productions for the prepared scenarios in two conditions; with and without the proposed system. When planning without the system, the participants visualized an actual stage in their head. The two scenarios used in this study were selected from scenario download service [12]. Both scenarios have nearly the same length and have already included not only the dialogue but also the stage direction, which indicates how to act out and reproduce theatrical productions of the effect of sound and lighting in actual performance. Before conducting the experiment, we removed the 12 stage directions from each scenario; 6 of them are about sound and the rests are about lighting. Each removed stage direction is absolutely necessary for each scene. For instance, one of the removed sound productions is "Chiming of a clock" which was put in the scene of "The clock informs 12 o'clock", and one of the removed lighting productions is "turn on global lighting" which was put in a diurnal scene. Without these theatrical productions, the stage situation becomes so unnatural. There-

fore, if participants visualize the theatrical stage clearly, they can notice the unnatural stage easily.

After these preparations, each participant conducted the task to plan for the two prepared scenarios with and without the proposed system. In order to give a fair evaluation, we chose 20 participants who don't have any experience in theater and the system. 10 of 20 conducted with the system first time, then conducted without it second time. The rests did in reverse order. The time limit for planning for one scenario was 20 minutes. Before the experiment, the participants were explained not to think the correction of the previously registered productions but to plan only the requisite production, which makes a scene unnatural if it is not put in the scene. We did not tell them about the number of removed productions, so they freely planned productions. With reading the scenario, each participant planned sound and lighting productions for the scene that they felt unnatural. In this study, each participant conducts the work to plan a production of both sound and lighting effect. For sound, they treated simple sound effects we prepared, such as the chime of a clock, the knock and so on. And for lighting, they thought about turning on and off two types of lighting; the global lighting and the spotlight. In the experiments, when planning with the proposed system, participants constantly reproduced the productions—the dolls' position, the effect of sound and lighting—while reading the scenario. By contraries, without the system, they planned by visualizing the theatrical stage in their head. They wrote down planned productions on a prepared paper. After all experiments, we measured the number of the correctly planned productions, which is concordant with the removed production.

5.1.1 Result

The result is shown as the Figure 8. Vertical line means the average number of the correct answers, and horizontal line means the case that separated by "for sound with the system (SWS)", "for sound without the system (SWOS)", "for lighting with the system (LWS)" and "for lighting without the system (LWOS)". The t-test was conducted to validate the difference between the correct answers with and without the system. As a result, we found the significant difference in sound ($P < 0.05$), and lighting ($P < 0.01$).

5.2 Questionnaire investigation

In certain science and technology exhibition, we conducted the questionnaire investigation to 30 participants who are interested in science and technology. After we explained them about the system schema and the system function, two or three participants used this system together for reproducing theatrical productions for a few minutes. Then, they answered the questionnaire—on a scale of one to ten— on the usability and the usefulness of the proposed system.

The result of this questionnaire is shown as Table 1.

5.3 Consideration

As shown in Figure 8, when planning with the system, the participants could notice more unnatural points of the sce-
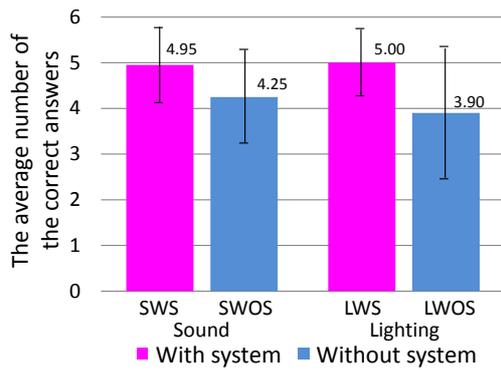
Figure 8: The average of the number of the corrected answers in two condition.

Table 1: Result of questionnaire evaluation.

Question	Point
Operation was easy.	8.6
Operation was intuitive.	8.9
Operation was joyful.	8.8
Workload was not arisen in simultaneous manipulation with other users.	8.1
Visualizing an actual theatrical stage was easy.	7.2
This system is useful for planning theatrical productions.	7.5
It is very useful to manipulate separately with respect to each role.	7.9

nario than without the system. We think this is because the system assists participants to perceive the theatrical stage as a whole by allowing to check the stage through the three-dimensional miniature theatrical stage reproduced on the table-top surface.

In the average number of the correct answers in planning lighting production, we found the more noticeable difference between with and without the system. This result shows the system allowing to actually reproduce theatrical productions is more effective for lighting production planning. For example, by actually reproducing and checking the theatrical stage situation through the miniature stage, the participants noticed the unnaturalness of *the pitch-dark stage* at "the scene that some actors/actresses are speaking in a house"; most people generally speak *in a well-lit room* when they are in a house. The participants also commented that the system for assisting to imagine the actual theatrical stage was more helpful in lighting production planning than in sound production planning. For these, we consider that the system contributes more useful support in lighting production planning, which is significantly associated with the stage situation.

In addition, in point of the standard deviation of the corrected answers, the value were 0.92 in SWS; 0.99 in SWOS; 0.84 in LWS; and 1.51 in LWOS. From these results, particularly in lighting production planning, we found the notice-

able difference between with and without the system. This is, as we think, because some of the participants could not keep on getting a clear picture of the theatrical stage without the system. For instance, without the system, they sometimes could not retain whether the global lighting was lighting-up or not. Like this, the users' ability to keep on getting a picture varies between individuals, thus the standard deviation became large. In contrast to this, with the proposed system, they could constantly look see the reproduced lighting production through the miniature stage reconstructed on the table, so they all could plan productions easily. As a result, by assisting users to visualize the theatrical stage as a whole, the system fills the gap between individual ability and contributes the work environment that whoever can easily plan a theatrical production.

Focusing on the system's contribution for multiple users' work, Table 1 indicates; *Workload was not arisen in simultaneous manipulation with other users* and *It is very useful to manipulate separately with respect to each role* received a high evaluation. From these result, the system function to contribute the personalized feedback to each user is helpful for users who have *the difficulty of sharing ideas of each production with other producers.* On this occasion, however, we only investigated how users feel about the system when using with others. In future study, we should investigate how the system affects multiple users' work from a quantitative standpoint.

For these, we can say the proposed system, which allows to reproduce theatrical productions on the three-dimensional miniature stage, is helpful for multiple users' collaborative work to plan a theatrical production by assisting users to perceive the entire theatrical stage.

6   Conclusion

Theater is the composite art which consists of various arts. In theatrical production planning, however, it is difficult to create a picture of the entire theatrical stage in their mind. To solve this problem, in this paper, we proposed a tabletop system for multiple users collaboratively working to plan a theatrical production. With this system, users can actually reproduce theatrical productions required for theater—the position of actors/actresses, the effect of sound and lighting—on a miniature theatrical stage reconstructed on the tabletop surface. Users can use the system function for reproducing theatrical productions by simple two operations; the one is to touch an operation window projected on the surface, and the other is to move dolls on the surface. In order to actualize the interaction with using physical dolls on the DiamondTouch table, we have developed the object recognition method that allows the system to detect the doll's type, position and direction when it is placed on the surface. In the evaluation experiments, we found that the proposed system helps users to plan a theatrical production by allowing to visualize the theatrical stage as a whole through the reconstructed miniature stage. From these, we can say the proposed system is useful for multiple users' work to plan a theatrical production.

Acknowledgments

REFERENCES

[1] Y.Nakamura and T.Hirano, The educational significance of the creation of drama -Through the analysis of the practice of integrated learning, "The project of creative drama"- (in Japanese), Bulletin of Tokyo Gakugei University, Division of Comprehensive Educational Science I, 61, pp. 57–71 (2010).

[2] M. Izima, Theater practice program vol.3 Lighting of theater stage vol.1 (in Japanese), Mondo Company (1997).

[3] P. Dietz, D. Leigh, DiamondTouch: A Multi-User Touch Technology, In Proc. UIST' 01, ACM, pp. 219–226 (2001).

[4] R. Oyama, Work of a professional in performing arts (in Japanese), Perikan Company (1997).

[5] H. Ito, The Staff: Work of the regisseur, Bansei Publishing, Inc., pp. 184 (1994).

[6] Q. Wu, P. Boulanger, M. Kazakevich, and R. Taylor, A real-time performance system for virtual theater, In proc. SMVC, ACM, pp. 3–8 (2010).

[7] M. Lewis, Bowen Virtual Theater, In Proc. SIGGRAPH' 03 on Web Graphics, ACM (2003).

[8] WYSIWYG. http://www.castlighting.com/.

[9] Matrix3.
http://www.meyersound.com/products/matrix3/.

[10] C. Dompierre, D. Laurendeau, Avatar: a virtual reality based tool for collaborative production of theater shows, In Proc. CRV' 06, IEEE, pp. 35–41 (2006).

[11] T. Takeuchi, K. Watanabe, T. Inoue, and K. Okada, DiamondTheater: A System for Reproducing Theater and Supporting Creative Activities, In Proc. HCI International' 09, Springer Berlin, pp. 416–425 (2009).

[12] Scenario Download Service.
http://www.haritora.net/script.cgi/.

Session 5:
Mobile/Pervasive Computing
(Chair: Yoshitaka Nakamura)

Development of a Historical Tour Support System supporting the Study Cycle

Satoru Fujii*, Maki Ushiroda*, Emiko Nanba*, and Kouji Yoshida**

*Matsue College of Technology, Japan
**Shonan Institute of Technology, Japan
*fujii@matsue-ct.ac.jp
**yoshidak@info.shonan-it.ac.jp

Abstract – We have developed a historical tour learning support system to study pre-tour and post-tour using a personal computer, and tour support using a mobile phone. Pre-tour, users can prepare using 3D graphics, a voice guide and Web resources accessed by personal computer. On tour, support is provided by GPS and voice guide via a smart phone and users can take notes and draw sketches using a touch pen. Post-tour, the system can use GPS data collected during the tour and 3D graphics to display the route taken by the tourist. In addition, users can check their understanding and recall by taking quizzes. About ten students tried this system and filled out an evaluation questionnaire. Most of them rated this system as useful and necessary for maximizing learning benefits.

Keywords: Smart phone, 3D graphics, Study cycle, GPS, Web

1 INTRODUCTION

Many historic sites provide autonomous sight seeing support systems. Almost all of these systems introduce sightseeing spots using the Internet and guidebooks. However there is benefit to the sightseer if the Internet information can be browsed by PC in the user's home before making a tour. Also there are few services to support tourist in real time at historic sites. Therefore the tourist should probably use only printed guidebooks or pamphlets. However, these printed materials can be slow to incorporate new information because the content cannot be updated frequently. First time visitors often cannot locate their current position or the part of the site they want to visit. Generally they take photos of historic sites, but they do not think of sketching it by hand.

"Shared Virtual Worlds for Mobile Phones" is one system using mobile phones and GPS [1]. This system displays 3D virtual city space and guides using a mobile phone with GPS. This system requires only a cellular phone without any new add-on devices. However, there is no function to enable users to record information themselves.

An alternative approach was the "Kyoto e! Project". This provided tourist navigation support using PDA（Personal Digital Assistant）[2]. This system had the function to provide information on the neighborhood using GPS data,

and a function to share users' information via the Internet. This approach was limited by the fact that this information was only accessible at specific locations. The "Learning Support System using Cellular Phone with GPS for Field Studies" could service proper contents for each user to get current position and his/her past action history [3]. This system was designed as a support to teaching materials, but had no personal record making function to share with another user.

"P-Tour: A Personal Navigation System with Travel Schedule Planning and Route Guidance Based on Schedule" could navigate a pre-set tour route using a mobile phone GPS function [4]. However, this system had no functions to guide by voice or Web page, or to record the tourist's route.

There are few systems that combine a study cycle, ability to get site-specific historic information in real time and personal tour record functions. These systems could not be supported ubiquitously, independent of the user's status. The paper "Experience Learning Support System in Integrated Study with use of Cellular Phone" adopted the study cycle that was constructed with pre study, study through experience and post study[5]. But this system did not use 3D graphics that was effective to understand leaning objects.

Therefore we are developing a historical tour learning support system modeled on Matsue Castle Park, which is our local historic site. Our system has the functions for pre tour using 3D graphics, tour using mobile phone and post tour using quiz [6][7][8].

The functions of pre tour are to learn about the historic site instinctively using 3D images. Having viewed the site in 3-D pre-tour, a tourist can appreciate the perspectives and important characteristics of the site when actually there. The role of pre tour study is to provide the root knowledge about what to look out for at tour time. Smart phone is used to make the file of tour track. The file is produced from GPS latitude and longitude records every ten seconds, and sent to the server when the termination button is selected. The file will be downloaded in the terminal for post tour learning. We developed the 3D virtual tour function for post-tour study. We wrestled with a method to re-create the user's walking trace in 3D space to enable review of his/her tour.

2 SYSTEM STRUCTURE

Figure 1 shows the structure of historical tour learning support system. This system supports three steps; pre tour, tour and post tour. Users can operate this system using a personal computer and smart phone. They can prepare before the tour using the 3D space virtual tour and WWW. While touring, the system provides guidance via a smart phone and also allows for photographs and note taking. Post tour, the user can re-trace his/her route in 3-D space and check memory and understanding via the quiz function. In this way, this system can increase learning effectiveness by supporting the study cycle; pre tour, on tour and post tour learning. It will be most effective to learn about historic site when user uses all of these modes according to the study cycle. But if user cannot use pre tour or post tour by PC, he/she can use only tour mode by smart phone.
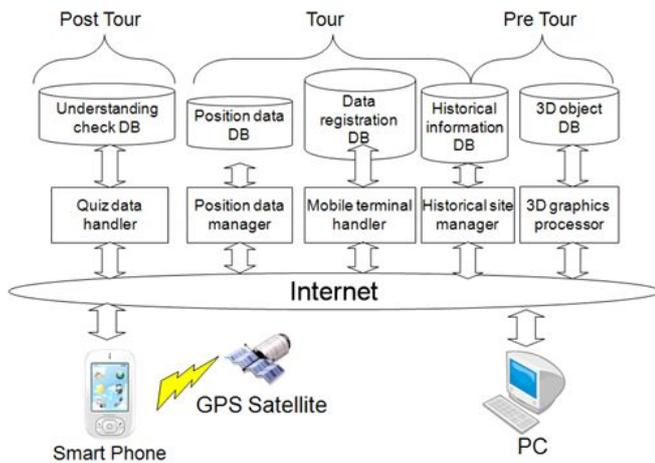
Figure 1 System structure

3 FUNCTIONS OF HISTORICAL TOUR LEARNING SUPPORT SYSTEM

3.1 Functions for pre tour

Pre tour, the user can study using the virtual tour in 3D space. Figure 2 shows the first screen the user would see after selecting this area in Matsue Castle Park. The user can learn about the historic site by moving the avatar using a mouse or keyboard.

The system has content for teaching the history, characteristics and layout of Matsue Castle, Matsue Local House, Gokoku Shrine and Jozan-Inari Shrine.

The user can move freely in virtual 3D space by clicking and dragging a mouse. Extended click action moves the avatar backward and forward and the view direction is changed up or down by dragging the mouse. Additional functions facilitate changes in walking speed and direction

of viewpoint. More detailed information is provided by the system voice guidance or can be got from Web pages, which are accessed by clicking on the historic site symbol in virtual space. The same voice data will be downloaded from the server using the smart phone GPS function when user actually visits the historic site.
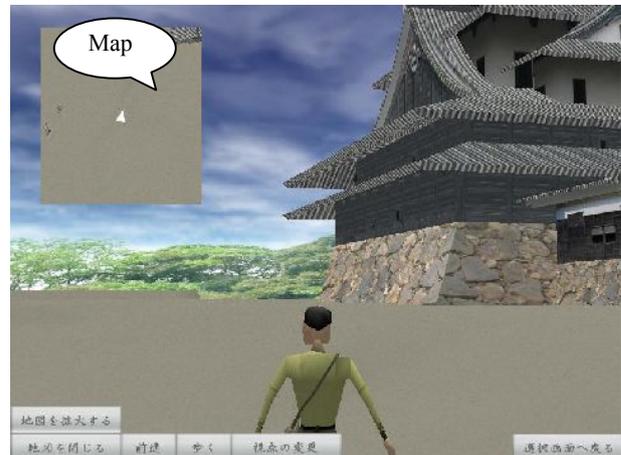
Figure 2 Sample screen of virtual tour

3.2 Functions for tour

3.2.1　Guide function

Currently, it is difficult for a newcomer to find the exact historic site information for their location from a pamphlet. Our system guides the user around the site according to the user's position by using GPS. Figure 3 shows the image of guide function.

Figure 3 Guide function

When the start button on the initial screen of the user's terminal is selected, a user name input form will be displayed. If the user name is input, the system will connect to the server and download the newest historic site

information to the user's terminal. At the same time, the system will get position data using GPS.

When the user activates the system and approaches the historic spot, the system will display information on lower portion of the initial screen. At the same time, the system will download guidance voice to the terminal and play it automatically. The "Replay guidance voice " button allows the user to listen again.

3.2.2 Historical material browse function

The browse function (Figure 4) accesses historical material, information viewed pre-tour and sketch data.

Figure 4 Historical material brows function

Historical material can be browsed by clicking on the "Information serviced by guide function" button. At present, the user can find many kinds of information independent of position. This information should be automatically updated so that the user has access to the newest information. Information available via the pre-tour function can be reviewed by clicking on the "Information used pre tour" button. By selecting the "Browse learning record" button, a user can browse the learning records he or she produced or those generated by another user. This function also allows records to be edited or removed. The "Sketch browsing" function allows users' sketches to be viewed. Clicking on the thumbnail image of the sketch expands the image. Tourists can study a historic site with reference to another user's learning record and sketches by communicating with the server.

3.2.3 Sketch function

Photography is forbidden in many castles and museums. Therefore, we developed a hand writing and sketch function to record images of historical material where photography is banned. Figure 5 shows a sample screen. Alternatively, where photography is allowed, the user can use this function

to annotate any photographs.

The pen for hand writing has six line thickness options and 48 colors. Since the screen of a smart phone is narrow, it is difficult to display all the color choices. Therefore only major colors are indicated on the main screen. All colors can be accessed by choosing the "another color" option allowing any of the 48 colors to be selected from another window The "Erase" button removes unwanted parts of the sketch and "Clear screen" button erases the whole screen. Colors used are recorded on upper screen, so previously selected colors can be used without opening the color palette. Sketched images can be stored on the smart phone and server as JPEG files. The images stored in the server can be browsed via "Sketch browsing" in the historical material browse function menu.

Figure 5 Sketch function

3.2.4 Creating a tour trace

One of the characteristics of this system was to be improved learning effectiveness by incorporating the data produced during the tour into the post tour learning function. To do this, we planned to produce tour route data, and send it to the post tour learning database. The tour route data file is produced from GPS latitude and longitude records every ten seconds. This file is sent to the server when the termination button is selected, and the file will be downloaded in the terminal for post tour learning. A user can review his/her tour track in 3D virtual space. Figure 6 is a sample of a trace on map derived from a set of tour track data. The traced curve is the route the tourist took.
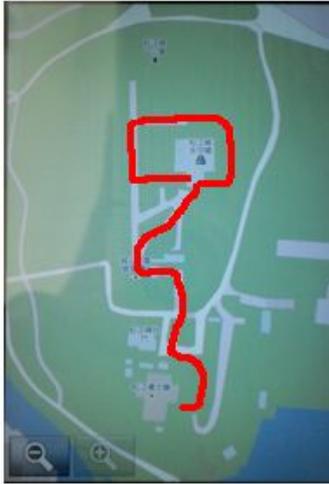
Figure 6 Trace of tour walking

4 FUNCTIONS FOR POST TOUR

4.1   Reproduction of tour trace

(a)   Surrounds of Matsue castle

(b)Around Matsue hometown mansion
Figure 7 Review of tour route scenes

It is necessary to make good use of reproduction in 3D space to use the trace data produced during the user's tour. The tour trace data is the latitudinal and longitudinal position information produced by GPS. This data is transferred to position data presented by plane coordinates to project it in 3D space. The next step is to smooth the transferred tour trace data to make the trace curve. Then the avatar can walk through on the trace curve, seeing everything from a camera position at the avatar's eye level.

Figure 7 shows a scene from the tour route review. The avatar walks through from Matsue Castle to Matsue hometown mansion in 3D space at real speed. Such a 3D space can be built using LightWave3D, a tool of 3D contents developing tool. In addition to this tool, 3D space is developed with Virtools, which builds with BB (Building blocks) [9].

The tour trace data is latitude and longitude from GPS, recorded every ten seconds during the tour. This data is downloaded to the personal computer from the server. The format of the trace data reflects that of latitude and longitude, do is shown as ddd.mm.ss.

The latitude and longitude have to be changed to plane coordinates of unit meters to reproduce the tour trace, because 3D space is expressed by the coordinates x, y and z. Change in position as meters per second, must also be calculated because the tour trace data changes are smaller than the second unit.
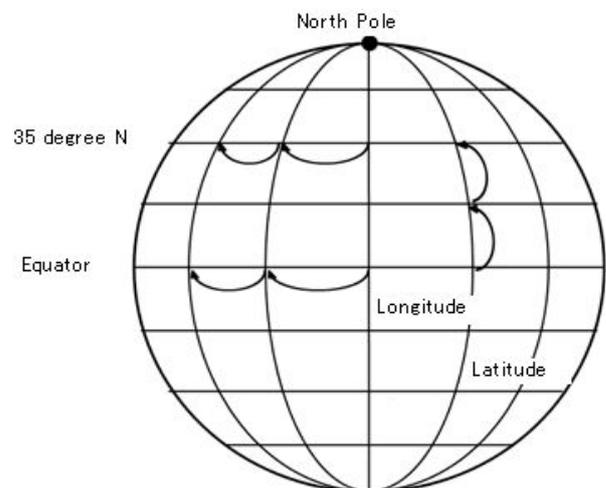
Figure 8 Difference in movement interval depends on latitude and longitude

The earth can be approximated by a sphere as in Figure 8. Latitude means lines drawn on earth parallel to the equator. Movement distance of latitude is solved from circumference of the earth. Circumference of the earth calculated by:
$2\pi R \fallingdotseq 40075020[m]$ (radius of the earth R＝6378137[m]).

If latitude changes 1 degree, distance is derived as in the formula.

$$2\pi R/360 ＝ 111319.5[m]$$

Therefore if latitude changes 1 second, distance follows:

$$2\pi R/(360*60*60) ≒ 30.92208[m]$$

Longitude is the line from the North Pole to the South Pole. When longitude is changed, the movement distance must be calculated differently. The movement distance is equal to latitude on equator, but radius cutting the earth along latitude changes less as latitude becomes greater, and becomes 0 on North Pole. Label "b" in Figure 9 that is the plane cutting latitude 35 degree N (the location of Matsue City). The radius r of the cut plane is calculated by the next formula if angle bOc is θ (see Figure 9).

$$r ＝ R\cos\theta$$

When longitude changes 1 degree, distance becomes:

$$2\pi r/360 ＝ 2\pi R\cos(\pi*35/180) ≒ 111313.2[m]$$

Therefore if longitude changes 1 second, distance is [10]:

$$2\pi r/(360*60*60) ≒ 30.92032[m]$$

By using this method to project in 3D space, the changes in position information can be expressed using planar coordinates. The transformed trace data is then used to make the tour trace curve. The avatar can walk through on the trace curve by programming BB of Virtools. The system can project the viewer's position during the tour by using a camera that has the avatar's eyes.
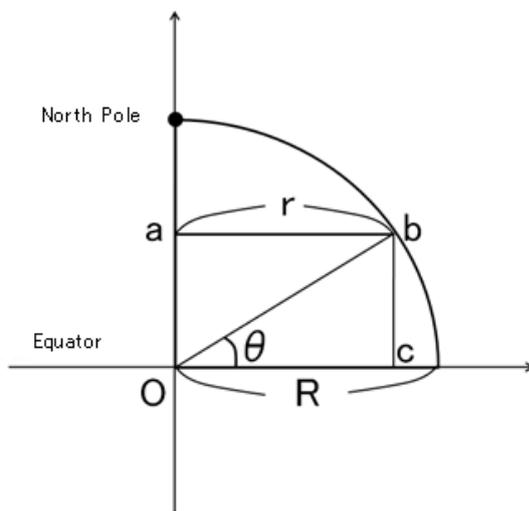
Figure 9 Movement distance when longitude changes

4.2 Understanding check

The understanding check quiz verifies how much the user understood about the historic site at pre-tour and tour. The quizes are constructed with elements that are elective questions and practical that asks filler questions. Each quiz is constructed with fifteen questions, therefore number of total question is thirty. Figure 10 shows a scene for quiz of filler question. If the user thinks about the answer, fills in blanks and selects "Answer" button, the system will respond correct/incorrect of each blank. Then the user can recognize about his/her understanding about the historic site.
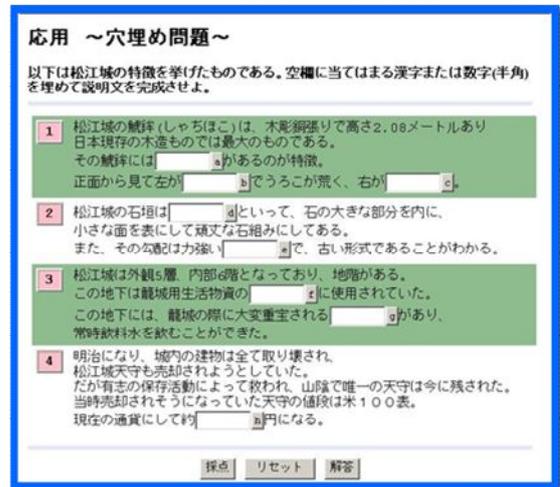
Figure 10 A scene of understanding check quiz

5 IMPLEMENTATION

3D graphics are developed by building BB using Virtools4.0, and C++ scripts are created automatically. Figure 11 is a sample screen of development using Virtools.

We used the 3D modeling software MetasequoiaLE R2.4 to produce the historic site objects. Metasequoia has an easy interface and plentiful polygon editor. We used LightWave Modeler 8.0 to post textures of historic site objects. If the Virtools plugin is added, 3D objects usable in Virtools can be created.
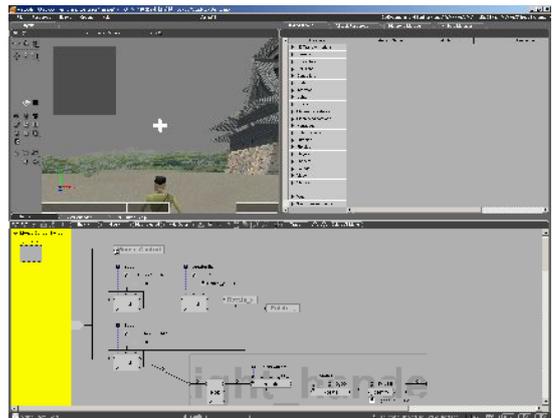
Figure 11 Sample screen of development using Virtools

The OS of smart phone is Windows Mobile6.1 Professional. Programs of smart phone are implemented with C# language using .NET FRAMEWORK.

6 EVALUATION

Twelve students tried the pre-tour learning functions and gave their evaluation by questionnaire. The questionnaire asked;

Q1. Could you operate the mouse smoothly?
Q2. Did you feel there were enough function buttons?
Q3. Could you move between different areas smoothly?
Q4. Was the quality of 3D objects, such as structures, good?
Q5. Didn't you feel somewhat out of place with the background of graphics?
Q6. Was the Web page useful for tour learning?
Q7. Was the voice guide useful for tour learning?
Q8. Did you feel the content of the quiz was good?
Q9. Do you think pre tour learning functions using 3D space are effective for preparing for your tour?
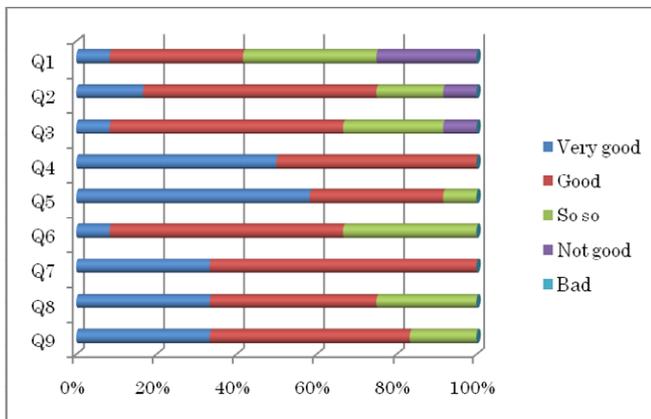
Figure 12 Evaluation of pre tour learning functions

Figure 12 shows the evaluation results. Q1 indicates that three users felt the operation with clicking the mouse was difficult and the operation manual was not easy to read. However, there was an opinion that operation by keyboard was comfortable. Results for Q4, Q6 and Q7 were hopeful for us. Some evaluators felt that Web pages were necessary for a tour, and that an on/off function was needed for the voice guidance. Evaluators felt that ground display was unnatural (Q5) and there were not enough quiz answers (Q8). Most users considered that the pre tour learning functions were effective.

Ten students tried the tour learning functions. They were subsequently asked:

Q1. Could you operate the system easily?
Q2. Did you feel comfortable carrying the mobile terminal?
Q3. Did you feel at ease operating the touch pen?
Q4. Could you use the guide functions smoothly?
Q5. Could you make your records smoothly?
Q6. Were the historic information pages useful?
Q7. Was the position display function useful?

Q8. Was the handwriting sketch screen easy to operate?
Q9. Would you want to use this system for your sightseeing?
Q10. Do you think any functions need improvement or development?
Q11. Please comment if you have another opinion.
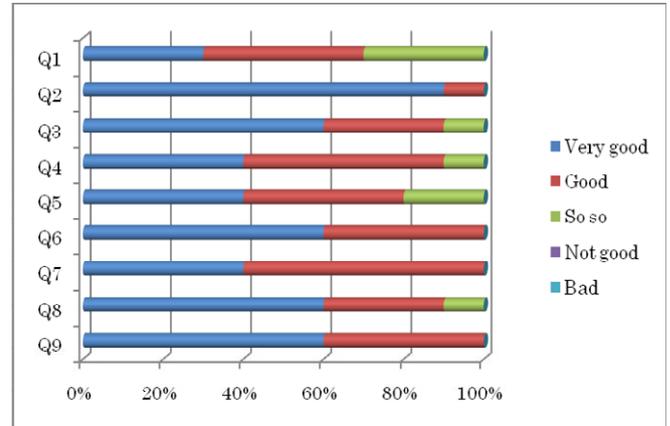
Figure 13 Evaluation of tour learning functions

Figure 13 shows the evaluation of tour learning functions. In response to Q9, all evaluators answered that they hoped to use this system on their tour. All thought that the mobile terminal was easy (Q2) and useful for getting historic site information (Q6). The responses to Q3 and Q4 were positive regarding ease of use of the sketch and guide functions.

The very positive evaluation for mobility is a result of using a smart phone rather than a PDA. The historic site information browse function elicited positive evaluation because its pages were constructed with images in addition to text. The convenience and learning effect responses had improved over evaluations of earlier systems due to the new additions of the voice guide and handwriting sketch functions.

Conversely, the relatively negative responses to Q1 and Q5, ease of operation and creation of learning records, show that these areas must be improved. We also need to think about adjustments in response to the opinions and remarks of the evaluators given in Q11.

Finally ten students tried out the post-tour learning functions and evaluated them by this questionnaire.

Q1. Did you feel the 3D virtual space looked like real space?
Q2. Was keyboard and mouse operation comfortable during the virtual tour?
Q3. Were you comfortable seeing the map and changing the viewpoint?
Q4. What are functions need to be improved or added for post tour learning?

Q5. Were the quiz contents suitable for review?

Q6. Were the quiz number and level measurable?

Q7. How about responding to the quizzes by keyboard?

Q8. What was your opinion of the avatar display by trace review function?

Q9. What are functions need improvement or what functions should be added for the route review?

Q10. Would you want to use this system for your historic site tour?

Q11．Please give your opinion or comment on your impressions.
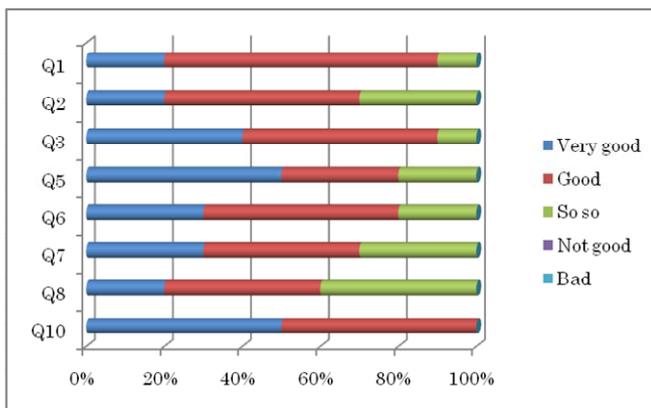
Figure 14 Evaluation of post tour learning functions

Figure 14 shows the evaluation of post-tour learning functions. Ninety percent of users answered that the 3D space looked like reality (Q1) and that the tour assistant functions were good (Q3). The contents of the quizzes were considered suitable for review by eighty percent of users (Q5) and all users answered that they hoped to use this system for their sightseeing (Q10). Five of the ten evaluators commented in their responses to question 11 that this system was interesting, delightful like TV game and impressive for structure's reality. This was a better result than we had hoped for.

However, only sixty percent of users were satisfied with the animation of the route reproduction (Q8). This was because there was an error of about five meters in the position data from GPS, and the trace curve could not be made correctly. The Japanese government has decided to launch stationary satellites in 2014 to construct Japanese GPS. If Japanese GPS were combined with American GPS, the error would decrease to less than one meter in the animated avatar display [11].

7 DISCUSSION

We considered the questionnaire results and identified the next set of challenges for the tour phase.

(1) Improvement of operation

It will be effective to review the design of smart phone screen, and add a "Return" button in each function to return the start page. It also needs a help function to guide beginners in the basic operation of smart phones.

(2) Improvement of function for making a learning record

Some evaluators found the create learning record function difficult to use effectively. So the screen design needs improvement and the "Free note" window should be reviewed. For example, this window should be divided into impressions of the historic site and advice for visiting there. Users should also be able to write information on these more materially.

(3) Replacement of the mobile terminal

We used a smart phone with a 4.1 inch screen as the mobile terminal. This screen size is too small so the characters and images on-screen are also too small. Also it is not easy to operate buttons on the screen. Therefore we are planning to upgrade to a tablet PC like GALAXY Tab of NTT docomo which has 7.0 inch wide screen. On the other hand, a tablet PC is bigger and heavier to carry than a smart phone. We need to find a balance between the priorities of ease of operation and ease of carrying.

The future challenges for post tour learning, as identified from the evaluations are listed below.

(1) Addition of 3D object

Evaluation of 3D was good, but it is necessary to further improve the reality factor. For instance we are planning to place objects such as trees. In addition, the passable route needs to be limited by putting in a transparent wall.

(2) Addition of a route selection function during re-trace review

This trial trace review used only one file created from GPS data. However the ability to select and display a file from re-trace route files is desirable. To improve this defect, menu screen should indicate the list of trace data available for selection by the user.

(3) Append voice guidance for trace review

Appending automatic voice replay as the user approaches the historic sire during post-tour review would create consistency with pre tour learning and tour learning. This historical tour learning system can support three phases; pre tour, tour and post tour learning. We evaluated each phase individually. As a next step, we must evaluate the system in combination, and identify the effects of the whole study cycle, which is the primary characteristic of the system.

8 CONCLUSIONS

We used 3D graphics and smart phone to create the Historical Tour Learning Support System to support tourists. We developed voice guide and handwriting/sketch functions to resolve problems identified in earlier support systems.

Ten users tried the system on tour and evaluated it by responding to a questionnaire. All of them hoped to use this system for sightseeing and evaluated each function as good. We developed tour trace function to record the tour route, and to incorporate tour learning with post tour learning. We believe that user can study more effectively by combining preparation and review.

For post tour learning, we considered a method to review the tour route and as a result, it is now possible for a tourist to follow the route they took in 3D space. Moreover we developed a quiz function to deepen and check understanding.

The ten students who tried and evaluated the system responded that the system was useful for historical tour learning, and most of them hoped to use this system at their historical tour.

We are planning to evaluate the system in combination. That is, to have evaluators undertake all three phases, pre-tour, tour and post-tour, to confirm its effectiveness. Finally we need to seek evaluation from the general tourist.

In future this system should be progressed to cover not only historical tour learning but also outdoors learning such as sightseeing, social visitation and school excursions.

REFERENCES

[1] K.Tarumi, Y.Tsurumi, K.Yokoo, S.Nishimoto, K.Matsubara, Y.Hayashi, Y.Harada, F.Kusunoki, Y.Mizukubo, M.Yoshida, S.Kim: Open Experiments of Sightseeing Support Systems with Shared Virtual Worlds for Mobile Phones, Journal of Information Processing Society of Japan, vol. 48, No.1, pp.110-124 (2007).

[2] Kyoto e! Project Information Politics of Ministry of Economy, Trade and Industry: http://www.meti.go.jp/policy/it_policy/rikatuyou/rikatuyo u_e-pro_6.htm (2007)

[3] Torino, M.Suzuki, F.Kushinoki, I.Hatano: Learning Support System using a Cellular Phone with GPS for Field Studies, The Institute of Electronics, Information and Communication Engineers, IEICE Technical Report, ET2007-88, pp.25-30 (2008).

[4] K.Matsumura, J.Kuroiwa, I.Takahashi, T.Okada, H.Ogura: Experience Learning Support System in Integrated Study with use of Cellular Phone, Transactions of Japanese Society for Information and Systems in Education, Vol.25, No.1, pp.7-17 (2008).

[5] S.Fujii, M.Ushiroda, E.Nanba: Development of a Historical Site Learning System supporting life cycle of study, JSiSE Research Report, vol.25, no.6, pp.3-10 (2011).

[6] A.Maruyama, N.Shibata, Y.Murata, K.Yasumoto, M.Itou: P-Tour: A Personal Navigation System with Travel Schedule Planning and Route Guidance Based on Schedule, Journal of Information Processing Society of Japan, Vol.45, No.12, pp.2678-2687 (2004).

[7] Satoru Fujii, Takahiro Shima, Megumi Takahashi and Kouji Yoshida: Development and Evaluation of a Historical Tour Support System using 3D graphics and Mobile terminal, KES2010, Part Ⅲ, LNAI6278, pp519-526 (2010).

[8] S.Fujii : Development and Evaluation of Ubiquitous Learning Support Systems Using Three-dimensional Graphics，eFest2008-Connected Learning, Presentation, Auckland (2008).

[9] Formula Virtools Homepage: http://www.3ds.com/products/3dvia/3dvia-virtools/ (2011).

[10] Latitude and longitude 1 second distance -Absolute Playing!-:http://d.hatena.ne.jp/Kiske/20090324/1237894846 (2009).

[11] YOMIURI ONLINE: http://www.yomiuri.co.jp/net/news/20110105-OYT8T00378.htm (2011).

A Study on the method of applying presence service to a lot of user

Yasuhiro Hashizume[*], Hironari Kobayashi[**] , Koji Yamamoto[*] and Osamu Takahashi[***]

[*]Mobile Communications Business Unit, NTC Corporation, Japan
[**]Research and Development Center, NTT DOCOMO INC, Japan
[***]Department of Media Architecture, FUTURE UNIVERSITY HAKODATE, Japan
hasizume@mob.ntc.co.jp

Abstract – Presence service in mobile networks is a subject for research and development as one of important elements for advanced network service platforms. The problem it has is to require user action to update its presence status. There are some new-generation UEs (User Equipments) that can automatically send presence information to its presence server. But, this is not applicable to all the UEs. This paper proposes a mechanism that enables controlling timing to get presence information from the Core Network (e.g. IMS).

Keywords: presence service, location information, IMS

1 INTRODUCTION

Mobile Network is evolving. With the introduction of LTE (Long Term Evolution)[1] services, mobile network is reaching to ALL-IP based network[2]. New multimedia services, such as IMS (IP Multimedia Subsystem)[3] based services, are being introduced into Mobile Network, which may not be based on network operator's services. Also, mobile devices are becoming equipped with Open Platform such as Android[4], which enable easy participation in new services with its capability for downloading applications.

In the macro scale of the network, many services of mobile network are moving to "Cloud." The information previously managed on a terminal or device is shifting to be managed on the Cloud. Terminals are now being replaced by smartphones. They usually have GPS (Global Positioning System) and WLAN (Wireless Local Area Network) enabled, which allows third party to provide services based on user's information such as location information.

There are very active researches and developments in services based on user's own information[5][6]. There are services that detect and use user's current location, services that provide best method for getting user's current location from GPS, Radio Access Network, and Wireless Local Area Network, and services that make use of combination of above measuring methods.

There is different kind of a service that uses other user's information in real-time. SNS (Social Network Service) makes it possible for a user to share his/her information with other users, but update of presence information usually requires manual operations. In some presence service, based on user's situational information, it updates user's presence information automatically. Although the service may be available to some users, the number of the applicable users is limited and thus it has not become a basic service, which manages and provides presence information of all the users in the network.

The expectation to have presence service for all users is high, if we considered the situation created by the Eastern Japan great earthquake that occurred in March 2011. At the time of a disaster outbreak, it is difficult for user to take effective actions. The network's limited resource should be used effectively, and it should be possible to collect presence information of all users.

In this paper, we propose a method that allows a mobile network (a carrier) to get current position (or presence information) of all users automatically and to manage it.

The problem to tackle is how mobile network can contribute in case of "outbreak (congestion) of enormous network traffic." Our approach is network (IMS) controls quantity of update of the presence information (traffic).

2 TARGET OF STUDY

In this chapter, we clarify network constitution and a target of the study.

2.1 Network constitution

The network is assumed to include an IMS based on several Access Networks. The mobile terminals (User Equipment) are assumed to have SIP (Session Initiation Protocol) capability. Access networks include GPRS (General Packet Radio Service), LTE, and WLAN, etc.

Terminals are assumed to be a mixture of various old and new ones. Presence service is assumed to be provided as a service function by the network (IMS), and not assumed to be installed as an application to the terminals.

2.2 Presence element

Presence information is information about the situation of a user, and it consists of several elements. The presence elements in 3GPP[7] are defined as follows (in Parlay X interface)[8][9].

Table 1: Presence elements defined in Parlay X

Element	Value
Activity	Busy, Steering, Lunch, Shopping, etc
Place Type	Home, Office, Hotel, Airport, etc
Privacy	Public, Private, Quiet, etc
Sphere	Work, Home, etc
Communication	Phone, Chat, SMS, Video, Web, etc
Mood	Angry, Cold, Happy, Lonely, etc
Place Is	Audio, Video, Text

In mobile network, "the present location" is the most important, and there are many studies of the actions to estimate by using a history of the positional information etc [10][11].

However, among elements of the presence information, "the present location" is the most frequently updated. In addition, it uses a lot of network resources to get the present location to carry out GPS and cell pinpointing operation processing. Therefore, it is necessary to prepare a vast network resource and a presence server to register and manage the present location of all the users with the regular network, which is difficult to realize.

In this study, we show a method of registering all the user's present locations on network (IMS) by implementing a function to judge the necessity of registration to a presence server.

2.3 Method of getting present location

There are several methods of getting a present location. One such example is use of Location Service (LCS) network service [12]. In this paper, however, the choice of methods is not a subject (There shall be constant network load independent of the method chosen). We are focused on the method of reducing the number of acquisitions of the present location information.

3 PROPOSAL METHOD

In this chapter, we present a method of updating the present location information of all users automatically in the network.

3.1 Difference of Stay and Move

The life of people can be divided into "Move" and "Stay." "Stay" means the person to be in the same place, such as in an office/company, in a class/school, etc. "Move" means the person is moving between the stays, such as currently moving by a train, a car, etc.

The information which has more practical value in presence information processing is "Stay," rather than "Move," which has high frequency of updates.

Therefore, in this paper, we implement a function to distinguish "Move" from "Stay" in network (IMS) and propose a method to make the network resource usage at the minimum level when acquiring the present location, by confining the acquisition of the present location to "a stay" only.
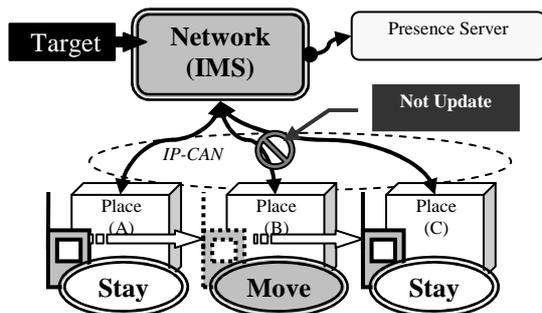
Figure 1: [Stay] and [Move]

3.2 Method to distinguish

Because a stay means a user with a terminal/device to be located in the same place for some time, we have implemented a function to watch staying time on the network (IMS).

When IMS detects that a user stays in the same area at constant time, it gets user's present location information (GPS information) and registers to the presence server. A watching timer is used to detect the user is staying or moving. When IMS detects the user is moving to other area while it is still watching the timer IMS judges that the user is on "Move", and IMS resets and restart the watching timer without updates the present location. While the watching timer is not up, the presence information is uniformly "Move", and IMS does not update.
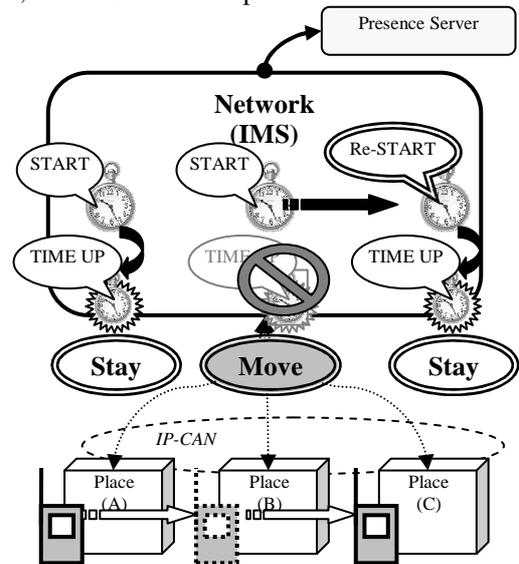
Figure 2: Method of judging

3.3 Problem of watching timer

A problem of this proposal method is that the present location registration to a presence server delays to watch a timer after the area movement detection. It is desirable to have as little delay as possible at time (i.e. watching time) in registering to a presence server from the time the area movement detection. But, in a case that the access area is large, if the watching timer is shorter than removable time in a cell, IMS may misjudge the movement for a stay.
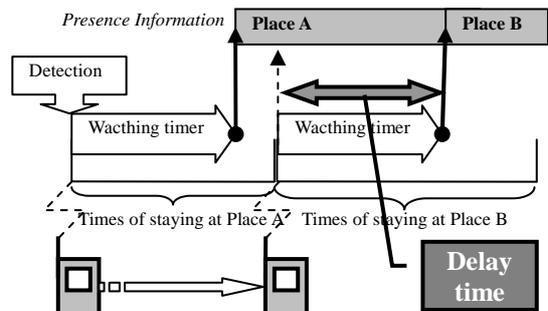
Figure 3: Problem of watching timer

As for the watching time, short period would be better if users think about information speed, but long period would be better from efficient use of network resources point of view. The reduction of network resource usage has a trade-off relation to the information delivery speed.

The setting of the watching time depends on the area size, and it is necessary to set the time accordingly. For example, if a diameter of the area size is 4km, a necessary time to cross the area by foot (as movement means to take the time most), e.g. one hour, is the watching time to be set up.

It is necessary to dynamically change the watching time according to the situation, to realize both reduction of the network resource usage and to provide the appropriate speed of the presence information delivery.

We propose two improvements here. The first is to change the watching time by a kind of IP Connection Access Network (IP-CAN). The second is to change the watching time by user's stay history.

3.4 Change of the timer by a kind of IP-CAN

IMS can include various IP-CANs. Depending on IP-CAN types, detection range of the area movement varies. It is, therefore, possible to make the watching time shorter by adjusting it, according to the kind of visiting IP-CAN.

The following table shows the characteristics of three types of IP-CAN: GPRS, wireless LAN (public), and wireless LAN (private).

Table 2: Presence elements defined in Parlay X

Type	Area Size	Stay Time	Probability of staying
GPRS	Large	Long	Middle
WLAN (Public)	Medium or Small	Medium or Short	Medium or Low
WLAN (Private)	Small	Middle	High

When an area is large, like the case with GPRS, it is necessary to set the watching time long. But, if an area is small and stay time is short, like the case with wireless LAN, it is possible to make the watching time shorter.

Therefore, the most suitable watching time should be set based on the characteristics of the IP-CAN, which is actually accessed when the user movement is detected in the area, assuming that IMS has full knowledge of and monitors the characteristics of various IP-CANs.
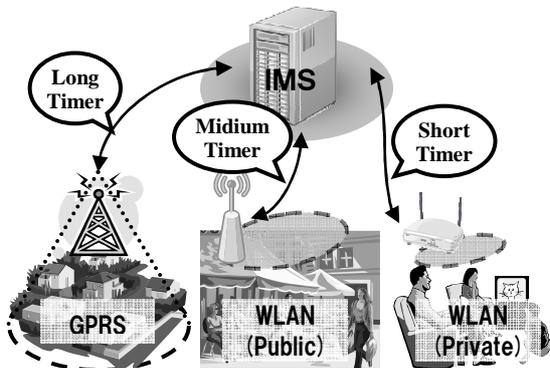
Figure 4: Timer value based on IP-CAN types

3.5 Change of the timer by a stay history

In real life, user may stay in the same place repeatedly, like at home, at the office/company, at shops. In other words, it is very likely that a destination area for a user to move to is often the area where the user stayed in the past.

Based on this observation, in some cases it is possible to make "Stay" decision earlier, i.e. making watching time shorter, even before watching timer completes, if IMS manages all of the user's stay history and a user movement is detected to the area that is found in the stay history.

This way, monitoring users' current locations with well managed history information including each user's visited areas will bring better serviceability of the network and efficient use of network resources.

3.6 Implementation method to IMS

Presence service standardized in 3GPP is used by this method. We show below five implementation methods to IMS [13].

1) Watching timer start at the time of the IMS registration
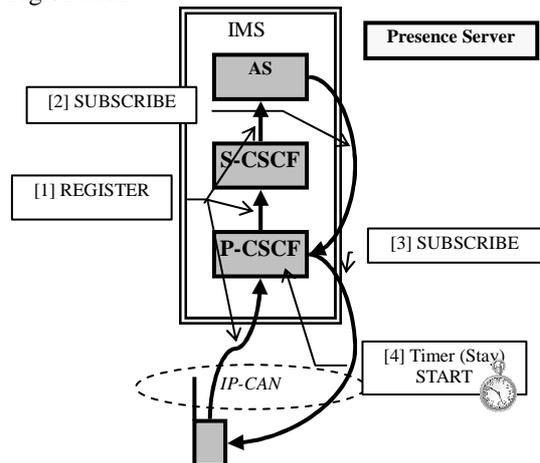
Figure 5: Architecture of watching stay time (Register)

In this case (Figure 5), a UE registers to Proxy-Call Session Control Function (P-CSCF), P-CSCF registers to S-CSCF, and S-CSCF registers to Application Server (AS) ([1]). The AS, then, sends subscription message to the P-CSCF to get updated Stay information ([2]). Next, the P-CSCF sends subscription message to the UE to get detected movement information ([3]), and starts the watching timer ([4]).

2) Movement detection before the watching timer expiration

In this case (Figure 6), a P-CSCF receives a message of movement from a UE ([5]) before the timer expires. It then reset the timer in action ([6]), and starts a new timer ([7]). If the resulting change is from Stay to Move, presence information should be updated as Move ([8]), without getting current location information. If the resulting change is from Move to Move, there is no need to update presence information. Only timer reset and restart will be done.
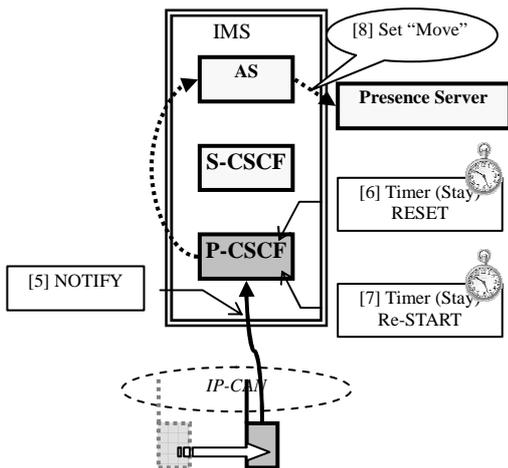
Figure 6: Architecture of watching stay time (Move)

If De-REGISTER is detected (e.g. power-off), P-CSCF will reset the timer. If it keep receiving movement messages, presence information will not be updated by assuming it is on Move.

3) Watching timer expiration (Stay detection)
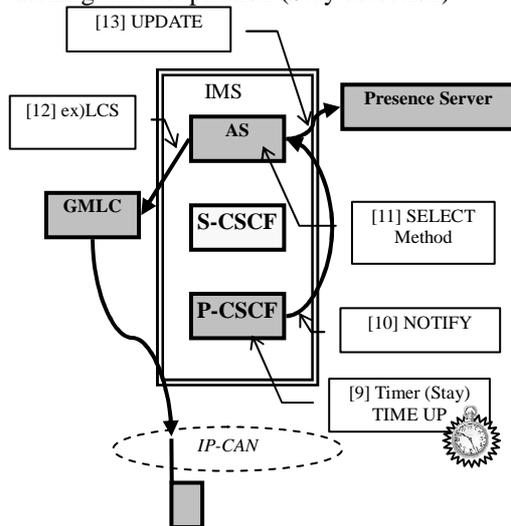
Figure 7: Architecture of watching stay time (Stay)

In this case (Figure 7), a P-CSCF detects Stay timer expires before it receives a notification message of movement from a UE ([9]). It then send notification message of Stay to an AS ([10]). The AS takes and executes one of the options to get user's current locations ([12]), and updates presence information ([13]).

4) Change the watching timer

In this case (Figure 8), a P-CSCF is notified with timer value and stay history. It starts timers with the timer value provided. It also compares user's current location information with the user's stay history, so that it makes watching time shorter if it found the current location in the stay history.
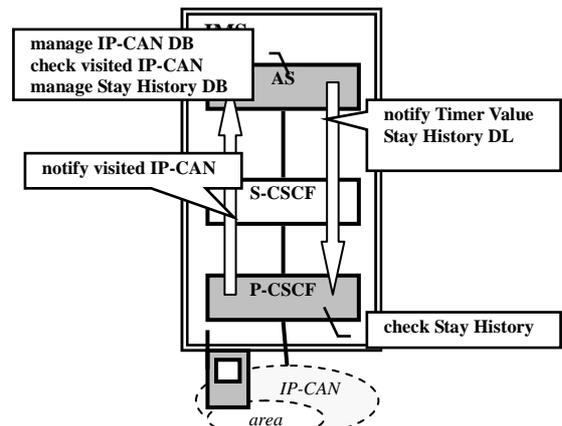
Figure 8: Function position
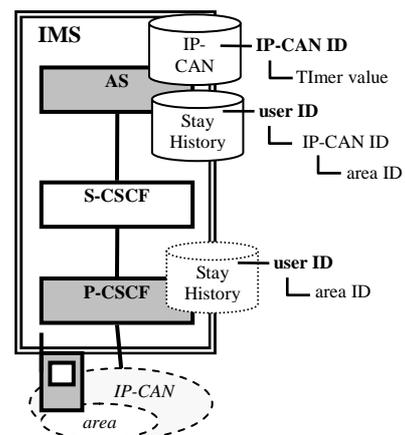
Figure 9: Data position

In this case (Figure 9), an AS calculates watching time value based on the visited IP-CAN's characteristics bound at REGISTER time using IP-CAN DB it manages. It also creates a list of location information of the user by retrieving the Stay History DB it manages with the corresponding IP-CAN identifier, and passes it to P-CSCF as additional information to subscription message to P-CSCF.

4   EVALUTATION BY SIMULATION

In this chapter, we show the simulation result of this proposal method.

4.1 Evaluation models

For an action of the general user, we use the following models.

1) The user commutes to the workplace which is 30km distant by a train (one time of transfer).
2) The user takes the lunch at a restaurant near the workplace (40 minutes stay).
3) The user goes on a business trip to the office of the customer once a day.

In addition, we simulate the following patterns of network constitution, at home, at station, at office of the user, at office of customer.

Table 3: Simulation patterns

Pattern	Home	Station	Office (user)	Office (Customer)	Restaurant
1	WLAN (private)	WLAN (public)	WLAN (private)	WLAN (private)	WLAN (public)
2	WLAN (private)	WLAN (public)	GPRS	GPRS	GPRS
3	WLAN (private)	GPRS	GPRS	GPRS	GPRS

In addition, it is assumed that the conditions of the network are as follows.

1) A cell size in GPRS is 2 km in radius.
2) The watching timer of GPRS is one hour.
3) The watching timer of public WLAN is 20 minutes.
4) The watching timer of private WLAN is 5 minutes.
5) When the location is included in the stay history, the watching timer is a half at decided time.

4.2 Simulation results

The result of the above-mentioned simulation patterns is described in Table 4. "Pattern" in the table is the same as "Pattern" in Table 3. "Traffic" shows the acquisition frequency of the present location a user a day. "Before" shows the numbers before the proposed method is used. "After" shows the result the numbers with the proposed method is applied. "Average" shows the average time (watching time) that IMS detects a user movement.

Table 4: Simulation results

Pattern	Traffic		Average
	Before	After	
1	37	7	4min
2	29	5	19min
3	19	5	19min

4.3 Consideration

The acquisition frequency of the present location has been reduced on about 80% average compared to the previous cases. And the watching timer has been reduced on about 77% (the watching timer before it improves was 60 minutes).

We judge the proposed method achieved very practical numbers from the viewpoint of reducing the acquisition frequency of the present location and from the viewpoint of shortening the watch time.

5 CONCLUSIONS

In the Eastern Japan great earthquake case that occurred this year, the traffic of voice-data communication for mobile phone increased rapidly for the safety confirmation of relatives. However, there were a lot of traffics, and the mobile phone network did not work as expected. Therefore, there were a lot of people who confirmed safety by using a SNS service that use the Internet connection. But, the situation/information that was able to be known in the SNS service was limited.

With this proposed method, it becomes possible for all users to manage the present location (includes "Move") always in the network. It is thought that the person who stays in home encounters some troubles when a disaster occurs. Oppositely, it is thought that the person who is moving has evaded from the trouble. Disaster and normal behavior are different. So the stay history in case of usual life may not use in disaster.

In the future, we will continue the research on the behavioral trait in case of natural disaster, and we will work on a method for more shortening of the watching time.

ACKNOWLEGMENT

REFERENCES

[1] 3GPP TS 23.401, General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access, Dec. 2010.

[2] Mikio Iwamura, etc, "Further development of LTE", NTT DOCOMO Technical Journal, VOL.18, NO.1, Apr.2010.

[3] 3GPP TS23.228, IP Multimedia Subsystem (IMS); Stage2, Dec. 2010.

[4] Android, http://www.android.com/

[5] Skyhook Wireless, http://www.skyhookwireless.com/howitworks/

[6] i-concier, NTT DOCOMO, http://www.nttdocomo.co.jp/service/customize/iconcier/

[7] 3GPP, http://www.3gpp.org/

[8] 3GPP TS29.199-1, Open Service Access (OSA); Parlay X Web Services; Part1: Common, Dec. 2009.

[9] 3GPP TS29.199-14, O pen Service Access (OSA); Parlay X Web Services; Part14: Presence, Dec. 2009

[10] Kengo Ohta, Masaki Mukai, Ismail Arai, Nobuhiko Nishio, Analysis of State Transition Generated from Observed Wi-Fi Access Points History, DICOMO 2010

[11] Masakatsu Aoki, Shunichi Seko, Masaaki Nishino, Tomohiro Yamada, Shinyo Muto, Masanobu Abe, Anestimating method for activity modes using location data, IPSJ SIG Technical Reports, Vol.108, No156, OIS2008-18, pp.7-12, July 2008.

[12] 3GPP TS23.271, Functional stage 2 description of Location Services (LCS), Dec. 2010.

[13] 3GPP TS 23.141, Presence Service Architecture and functional description, Dec. 2009.

A STUDY ON CONSTRUCTION METHOD OF SUBSCRIBER DATABASE FOR MOBILE NETWORK

Tatsuo Matsukawa,Nobuo Takeuchi,Tohru Hasegawa and Hironari Kobayashi

Core Network Development Department,Research and Development Center,NTT DOCOMO, Inc., Japan
{matsukawat,takeuchin,tohru,kobayashihiron}@nttdocomo.co.jp

Abstract -Mobile terminals of new forms of usage like smart phones, pad type terminals, and communication modules, etc. are increasing rapidly, and the scale out of HLR/HSS that manages subscriber information is required to be done flexibly and promptly according to the number of subscribers, the traffic, and the service, etc.
In this paper, we propose the method of separating the call processing part and the subscriber database part as one of the ideas to achieve the requirement. We also discuss the construction process of this system, some of the problems to be solved, and the approaches that can be taken to solve the problems.

Keywords: HLR, HSS, Scale out, Call Processing function, Subscription Database, Separation

1 INTRODUCTION

In the mobile communication network, the subscriber data, such as subscriber's location information and their contract information are maintained by HLR (Home Location Register) / HSS (Home Subscriber Server).

The subscriber data is referred, and updated at the time of location registration, call control (such as originating and terminating call), or when the subscriber's service information is changed at the mobile phone shop etc.

The data maintained in HLR/HSS increases if any new service is provided or, the number of subscribers increases. Especially, in recent years, the traffic characteristic (such as, location registration and originating and terminating call frequency), and the subscriber information maintained in HLR/HSS are diversified more and more, as the number of smart phones, and machine communication services (communication with the communication module terminals built into vending machines and cars, etc.) are increasing along with the development of LTE services. Moreover, the subscriber data management system using the phone number belt/band is also becoming difficult because of the influence of MNP (Mobile Number Portability).

This paper describes the problems of existing HLR/HSS and proposes the solution methods.

2 THE REQUIREMENTS FOR SUBSCRIBER DATABASE FOR MOBILE COMMUNICATION

The requirements of subscriber data base for mobile communication are as follows.
a)   Consistency
b)   Fast reference and update of data (minimization of delay caused by TAT (Turn Around Time))
c)   Continuation of operation

Though the data access, for example, Service Order(SO), location registration, or data access from other subscribers occur asynchronitically, it is necessary for the subscriber database to guarantee the order of access and to prevent any kind of state-unmatching compared to the intention of the user due to logical contradictions, even if any trouble etc. occurs. Moreover not only the references but also the updates occur frequently in the subscriber database for the mobile communication. The necessary amount of data for one subscriber that should be maintained or, the volume of data exchanged per database access is not too large. It is necessary to design the database that considers the above-mentioned features. Moreover, needless to say, as it can be assumed that the service is available for 24 hours and 365 days, it is necessary to continue serving by using the reserve system in the case of system failure or, during the operation of function addition, subscriber data transfer, etc.

3 THE PRESENT METHOD AND THE ISSUE

To construct HLR/HSS that met the requirements described in the preceding chapter, the method of accommodating call control function and database function in the same processor was adopted in existing HLR/HSS in consideration of the hardware specs etc. at that time.

Therefore, when user's traffic increases to rise the CPU utilization of the processor or, the volume of data that can be accommodated is exceeded (even if the CPU utilization is not high), it was necessary to add processors, to reallocate the range of accommodated subscriber number range and to transfer data to other processors in every case (Figure 1).

And because the nodes like LMSC (Local Mobile Switching Center) etc. that access to different HLR/HSS by the subscriber's number accommodated by a particular HLR/HSS and decide to which HLR/HSS they should access, the update process of the data related to the accommodation destination was done at each data transfer and the maintenance operation was required.

Furthermore, as the same processor to handle call control function and database function, memory allocation and control   timing must be optimized at each case. Therefore, when data elements are added, the program files are needed to be updated and it is difficult to add data elements promptly.
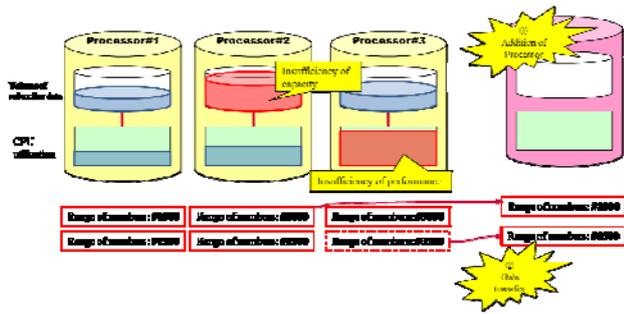
Figure 1:Scale out of present method.

4 SEPARATION OF CALL CONTROL FUNCTION AND DATABASE FUNCTION

Tables As an idea to solve these issues of the exsisting system described in the preceding chapter, the Front-End function (FE) that processes call control from LMSC etc. and the database function (DB) are separated. FE does not maintain the subscriber data and accesses to DB only if necessary [1]. Because it can solve the accommodated location of the subscriber data in FE and DB, LMSC can access to any FE without considering the accommodation range of subscriber numbers, and it is possible to get the same result.

Thus by separating FE and DB, the scale out of HLR/HSS can be made flexible. For example, when the addition of CPU resource is required because of a traffic increase, it only has to add FE. Similarly, when the increase of the subscriber data capacity is required, it only has to add DB.

Moreover, in the past method, though it was necessary to update routing information of the node connected with HLR/HSS when the accommodation range of subscriber numbers is change, these update work will not be required in the proposed method by solving the accommodation in FE and DB (Figure 2).

Furthermore, as the function addition to the call control and the element addition of DB can be made independently, a prompt service introduction and the reduction in the development cost can be expected.
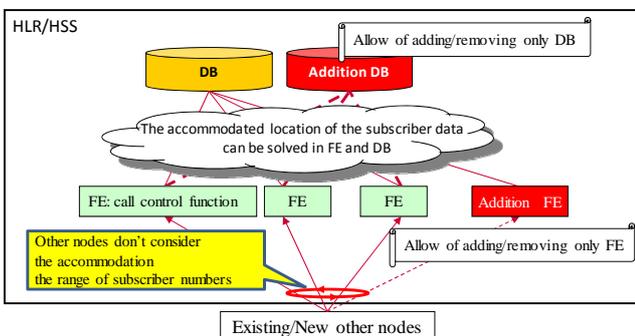
Figure 2:Separation of FE and DB.

5 CONSTRUCTION METHOD OF DB FOR SEPARATION FE AND DB

It is necessary to distribute data to two or more processors and storage disks (ex. HDD) because the number of subscriber numbers accommodated in HLR/HSS is too large (50 million or more in case of NTT DOCOMO) to process it with one processor. At the present time, it is thought to be difficult to meet the requirement for the performance of DB in construction with DB based on the disk (ex. HDD) due to the problem of the I/O performance.

And for the separation of FE and DB, though there is a concern that the processing response to the demand by the node connected with HLR/HSS will be late, the requirement level can be cleared with the improvement of CPU and memory hardware performance, the adoption of high-speed and large-capacity transmission.

6 ISSUES OF THE USE OF SUBSCRIBER DATABASE BASED ON MEMORY

In case of the use of subscriber database based on memory, because there in the amount of the memory that can be deployed in one processor, the volume of data that can be managed with one processor decreases considerably compared with the DB based on disk method. Therefore, it is necessary to distribute the accommodation processor of each some range of numbers (for instance, one million numbers) and to allocate it. Moreover, because data progressed to the memory is cleared at the processor failure, it is necessary that the memory and the processor are completely redundant or data in the memory is backed up to the disk etc.

When the processors are redundant completely, in normal circumstances, though standby system device might not be operated so much, it must be always operated for the same number as act system, and there was a problem of high cost from the equipment construction and the operation (See Fig. 3). Accordingly, instead of completely redundant system, it is preferable to be able to operate it by the necessary minimum number of the processor that alternative operation can be done in consideration of the failure frequency (N-ACT composition). To enable such a composition, the method that data is backed up to the disk and the processor is allocated dynamically if necessary, is adopted. In this case, there is a problem that data returns until the point backed up to recovery from the backup data to the memory. Therefore, it is preferable to be able to back up data in the memory as in real time as possible.
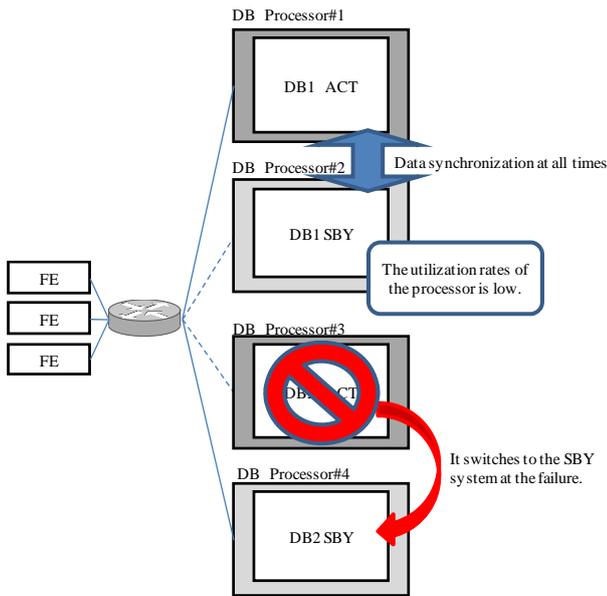
Figure 3:Redundancy of processor.

7 DATA GUARANTEE OF DB

It is preferable to be able to back up all data in the memory to the disk, considering restoring data. But it is difficult, considering the reading and writing processing performance of the RAID device at present. Therefore, it is effective to manage the data guarantee separately for the data in real time backed up, the data periodically backed up, and the data not backed up according to the characteristic of data.

The data maintained with HLR/HSS of DOCOMO is classified when roughly dividing as shown in the following Tables 1.

Table 1:Data classification of data in DB.

Information classification	Example of elements	Influence in the processor failure and restoration	Backup cycle policy
Service Order(SO)	Contract information	A state different from an actual contract is generated, and the account problem etc. occur.	Real-time
Subscriber's setting	Starting time of voice mail service	A real-time backup is hoped for as much as possible because it becomes different from the intention of the user setting.	Cycle buck up and difference (at once)
Network control	Maintenance information	There is no influence in service, and if maintenance person sets it again it is possible to restore though it becomes impossible to examine the function confirmation.	No buck up

It is the basic policy that the priority of the backup of the data that can be restored by means is lowered, and the one that influences to the charge and cannot be restored is higher. We will examine the classification of the data element in detail in consideration of the influence.

8 SHORTENING THE DATA RECOVERY FROM DB

There is a problem that the service interruption time becomes long when the volume of recovery data from the backup to the memory is large (See Fig. 4). As measures of this problem, data is divided by the constancy amount

(range of subscriber numbers), and of each is stored in another partition, when the trouble occurs in the accommodated processor, shortening the service interruption time becomes possible because it distributes to two or more processors that can be other available one and the data of each partition is recovered (See Fig. 5).

On the other hands, though the recovery time of each partition can be shortened by making the size of the partition detailed, the time of the partition mount to increase because the number of amount of partitions increases and to take time to the internal routing switch instruction, the service interruption time as the total is longer. It is necessary to examine the best size of the partition to which the service interruption time becomes small in consideration of the number of processors and the accommodation volume of data.
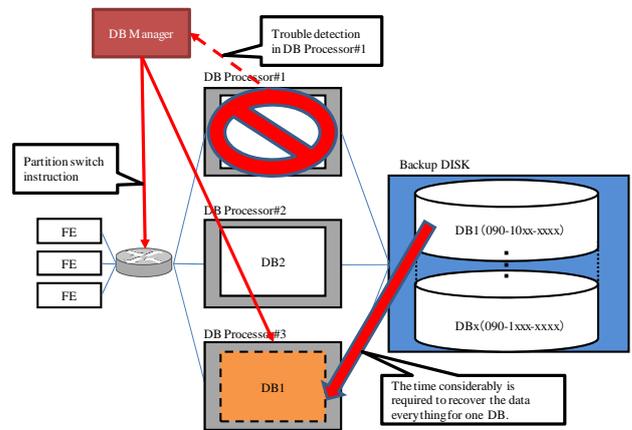
Figure 4:Recovery data from the backup to memory in failure DB.
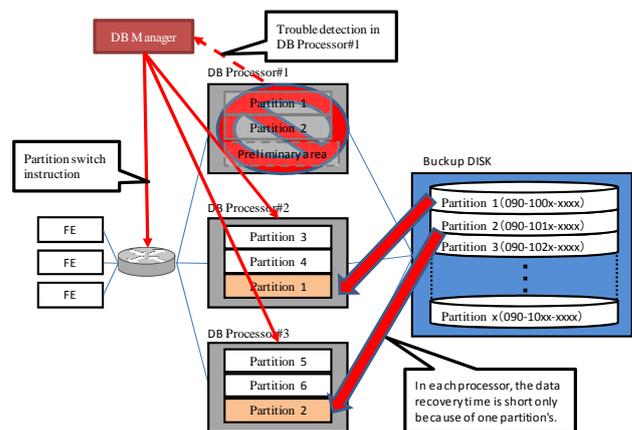
Figure 5:Shorting of recovery by partition division.

9 DATA RELOCATION

The idea into which data was divided by the partition in the preceding chapter was presented. By allocating the partition accommodated by the processor with higher loads to the processor with lower loads, the leveling of the load and accommodation can be aimed at, and equipment be used

efficiently. It is necessary to examine the best size of the partition to which the service interruption time becomes small in consideration of the number of processors and the accommodation volume of data. It is necessary to always manage the resource usage condition in the unit of the partition and each processor to execute this ' relocation' according to appropriate timing. Moreover, when a physical resource is insufficient even if it levels it, the scale-put of DB is possible by doing the relocation to the additional processor without the service interruption (See Fig.6) [2].
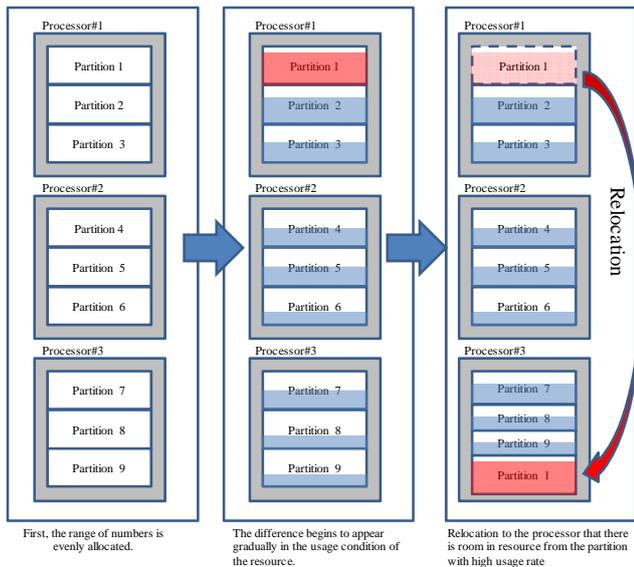
Figure 6:Image of relocation.

10 CONCLUSION

In this paper, we described the construction process of the Database in case of the separation call control function and subscriber database function, and advantage of it. And the settlement plan by the relocation was shown as shortening the service interruption time in the recovery of the backup data, the resource usage condition and the scale out process, in case of adopting the database based on memory method.

We will continuously examine the introduction of the method of the database separation that the new service introduction is enabled early, as soon as possible.

REFERENCES

[1]Tomonori Kagi et al., 2010, The separation of call processing and database function in IPSCP, Proceedings of The 2010 IEICE Society Conference. Osaka, Japan.

[2]Keiichirou Otsuka et al., 2010, Construction method of subscriber database, Proceedings of The 2010 IEICE Society Conference. Osaka, Japan.

Proposal and Implementation of Pseudo Push Using Network Subsystem and Task Execution for PC

Kazuaki Nimura[*], Hidenobu Ito[*], Yousuke Nakamura[*], Akira Shiba[*], Nobutsugu Fujino[*]

[*]Fujitsu Laboratories Ltd
4-1-1, Kamikodanaka, Nakahara-ku, Kawasaki, Japan
{ kazuaki.nimura, itou.hidenobu, nkmr, shiba.akira, fujino }@ jp.fujitsu.com

Abstract –For the progress of Cloud computing, receiving a service from the Cloud 24/7 will improve the experience of PCs (personal computers) like that of smartphones today. At the same time, an issue of electrical power comes to attention widely and increasingly. By putting a PC in lower power mode more actively, it can increase the sleep mode timing and can improve the electrical issue. However, this should not bring any degradation of the usability; otherwise it cannot be accustomed user to the policy. In this paper, we propose a system that can send a task from the Cloud and execute the task at a PC whenever it is necessary, even though the PC is in sleep mode. Therefore, it is easy to live with the sleep state of PC because it will open up the task automatically, if necessary. Hence, we prototypes the system that use a web application as a task and a pseudo push as a notification. Then, we evaluate it in a real field. As a result, the proposed system is found to be capable of executing the task in a timely manner. In addition, we confirm that the pseudo push can maintain the communication using little additional power consumption.

Keywords: PC, Push, HTML5, Application cache, Power consumption.

1 INTRODUCTION

Like a smartphone, it is desirable that PC remains connected to the network and receive services at all times. To receive network services, the PC needs to be powered on even when user does not use it, such as during a meeting or at nighttime [1]. To reduce the power consumption more actively, a tool is provided by PC companies, which can set the power policy easily and put the PC in sleep mode for a short period when there is no activity. However, actively putting the PC in sleep mode may have an impact on business performance because it cannot access network among them. In addition to put the PC in sleep mode more actively, any degradation of user experience is not desired.

In this paper, we propose an architecture that can receive and execute a task whenever needed. The architecture consists of four functions: task provision, task notification, task preparation, and task execution. Task provision and task notification are functions included in the Cloud side. Task notification is realized with pseudo push by introducing a subsystem that works even while the PC is in sleep mode. It can help in informing a user regarding the incoming task from the Cloud.

Then, we prototype the system on the basis of the usage that smartphone sends a task which consists of a photo viewer programmed as a Web application and photo data. Here, we focus on the Web application because the Cloud technology is widely used recently and many applications are provided as Web applications, such as HTML5 (HyperText Markup Language version 5) [2]. By using HTML5 as an application, it can provide a highly functional Web application that is comparable with the native application and can provide the functionality of a locally executable task. As the execution environment, we use the Google Chrome browser [3] and Chrome OS [4] because of its improved compatibility with HTML5.

Moreover, we evaluate it in real field environment. The result shows that the PC subsystem can deal with pseudo push, even when the PC host is in sleep mode, by using an incoming message. In addition, it can provide low-power downloading and execute the application without user access. Although we introduced an additional dedicated hardware to receive services, it proves that it is effective in power consumption.

The remainder of this paper is organized as follows. In section 2, we summarize related works. In section 3, we present our architecture. In section 4, we show the details of the prototype implementation. In section 5, we show the results of the system evaluation, and finally in section 6, we state our conclusions.

2 RELATED WORKS

In this section, we summarize the related works from the aspect of wake function as described in Table 1.

To wake through a LAN (Local Area Network), a magic packet needs to be issued as an awakener, like Local server. The controllable area is basically the local area.
- An architecture that can wake PC and continue network services for PC is described in [5]. This work copies the host system's properties to the subsystem in Sleep mode.

To wake through a WLAN (Wireless LAN), the PC needs to ask a proxy or an AP (Access Point) to issue a magic packet. The controllable area is basically limited in the local area.
- Apple Inc. has implemented proxying as a wake on demand feature on their wireless network-attached storage devices and computers [6].

To wake through a WWAN (Wireless Wide Area Network), an SMS (Short Message Service) gateway or a push gateway is used. It is needed to deal AT command to receive the message.
- About notification service, Intel Anti-Theft uses SMS as the notification message [7].

・Android Cloud to Device Messaging Framework [8] for Android phones and tablets use persistent connection by TCP (Transmission Control Protocol) and wake through the push gateway.
・How to establish a WWAN connection from a host system even when the subsystem has a network connection is described in [9].

Table 1: Types of network and wake for PC

Network type	Controllable area	Awakener
LAN	Local	Local server
WLAN	Local	Independent proxy, Access Point
WWAN	Wide area	SMS gateway

As far as we know, no feasibility study has been made on PC wake using the push gateway. The concept is briefly explained in this paper, however the main focus of our research is dealing with the tasks.

3 PROPOSED ARCHITECTURE

In this section, first we will explain the issue, and then show the requirement to solve the issue, in addition propose architecture to meet the requirements.

3.1 Problem

The problem we are focusing on in this study is there is no mechanism that can send a task whenever needed while keeping PC in a low-power state as long as possible. The term "task" in this paper means small job such as registration to an event through a network.

3.2 Requirements

The requirements to solve the problem statement are as follows.
(i) To have a function that can accept and send a task from the Cloud whenever necessary.
This corresponds to the phrase of "send a task whenever needed" in the problem statement.
(ii) To have a function that can notify a task to PC regardless of the power state of the PC.
This corresponds to the phrase of "while keeping PC in a low-power state as long as possible" in the problem statement because if there is notify capability, PC can be in sleep mode.
(iii) To have a function that can prepare the task without waking up the PC.
Usually a powered on PC consumes much power and if the task can be prepared without waking up the PC, it will contribute to reducing power consumption.
This requirement corresponds to the same statement described in (ii).
(iv) To have a function to execute a task without user interaction and show the task to user.

This corresponds to the same statement described in (i) because if a task does not execute in PC the cloud cannot receive the acknowledgement of "send a task". By complying with these, a task can be transferred to PC anytime and the problem can be solved.

3.3 Architecture

To meet each requirement, we propose an architecture that consists of following functions, which are also shown in Figure 1.
(I) Task provision: To fulfill requirement (i), a function that can provision a task, which will be transmitted to the PC, is introduced. As described later, a requester registers a task to this function.
(II) Task notification: To fulfill requirement (ii), a function that can notify task existence by messaging even if the PC is in sleep mode is introduced. For this, a network connection needs to be available anytime. Depending on the usage of PC, an adequate communication technology changes. If a PC is required only local network connection for example because of the large size, a communication technology which covers local area can be used. On the other hand, if we expect mobility to PC, wide coverage of the communication needs to be hired.
(III) Task preparation: To fulfill requirement (iii), a function that can do background preparation even if the PC is in sleep mode to contribute to reduction in power consumption is introduced. There are two type of background service. One is a task which is needed to wake PC, and consumed the task by PC. Another is a task which is not needed to wake PC, and it can transact all in background.
(IV) Task execution: To fulfill requirement (iv), a function that can provide and execute a task that is received even if the PC is in sleep mode is introduced. To translate the incoming task to executable task takes important role. Once PC turns on, a task is executed automatically or showed the existence to display.

By using those, PC can receive a task whenever from the Cloud. Therefore even persons who are inexperienced in PC can receive a service easily. Then, it can contribute to improve the user experience of PC.
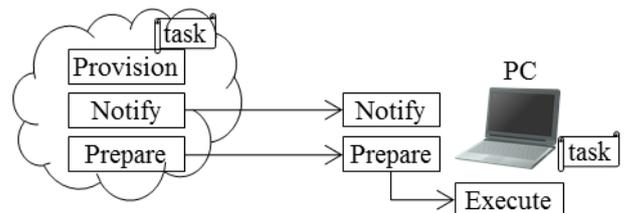
Figure 1: Proposed architecture.

4 PROTOTYPE IMPLEMENTATION

In this section, we will explain the prototype that realizes the proposed architecture and the task.

4.1 Prototype

Figure 2 shows the system structure of the prototype. The proposed system consists of the following hardware components: a requestor, task server, push gateway, host PC, and subsystem.

Requester:

There are two requesters. One is a requester which registers a task to task server. Another is a requester which requests a push to push gateway to send the task to the PC with the subsystem. It might be the requester is a corporate IT for example. This component corresponds to the function (I) in the architecture.

Task server:

The task server carries out (I) task provision. It receives a task from the requestor, adds it to the repository, and exposes the repository included tasks to the Cloud.

Push gateway:

The push gateway carries out (II) task notification. It is designed that it can provide the service not only PC but also other devices such as smart phone and smart tablet. It consists of the following three elements that asynchronously work with each other.

・ Push reception: A push message request is received from a pre-defined device such as smartphone, PC, or server.

・ Push sender: It is responsible for sending a push message to the PC. It manages a persistent connection with the PC to deal pseudo push. If there is no connection between the push gateway and PC for some reason, then it will resend the push message later.

・ DB: It has information of PC identifications and requester identifications. When it receives push request with a PC identification and requestor identification, the DB is checked whether the requester is preregistered and the PC identification is preregistered. If there are, it will serve.

Host PC:

The host PC cares the (IV) task execution and it has the following elements:

・ Browser synchronizer: It checks whether the web application or data exist in the SD memory and if there correct through the memory. The timing of check is at the return from sleep or polling from the host system while it is in the power on state. A packaged application or data are reconstructed by this function. In addition, it injects the application or data to the browser cache. By doing this, the Web application or data can be used after the browser is initiated or reloaded.

Subsystem:

The PC subsystem carries out the counterpart of the (II) task notification and the (III) task preparation and it consists of the following elements:

・ Push handler: It establishes a network connection with the push gateway. And it receives a task notification. In addition, it analyzes an incoming push message. If it downloads or PC control requests, then the following functions are called.

・ Downloader: If a push message has a download request, then it will download the Web application or data from the task server through the network to SD memory.

・ PC control: If a push message has a PC control request, then it will issue a wake command to the PC with a condition. A condition might be nothing or wait until something happens. The example of a condition is whether the user is in front of the PC or not. If not, the request of wake will be pending, and if the user comes to the PC, then it will issue a wake to PC.

As mentioned above, these components can suffice the four pillars of the architecture. Therefore by realizing the prototype, it can put the architecture into shape.

4.2 Task and functions to deal task

To express a task, we use a HTML5 application and/or
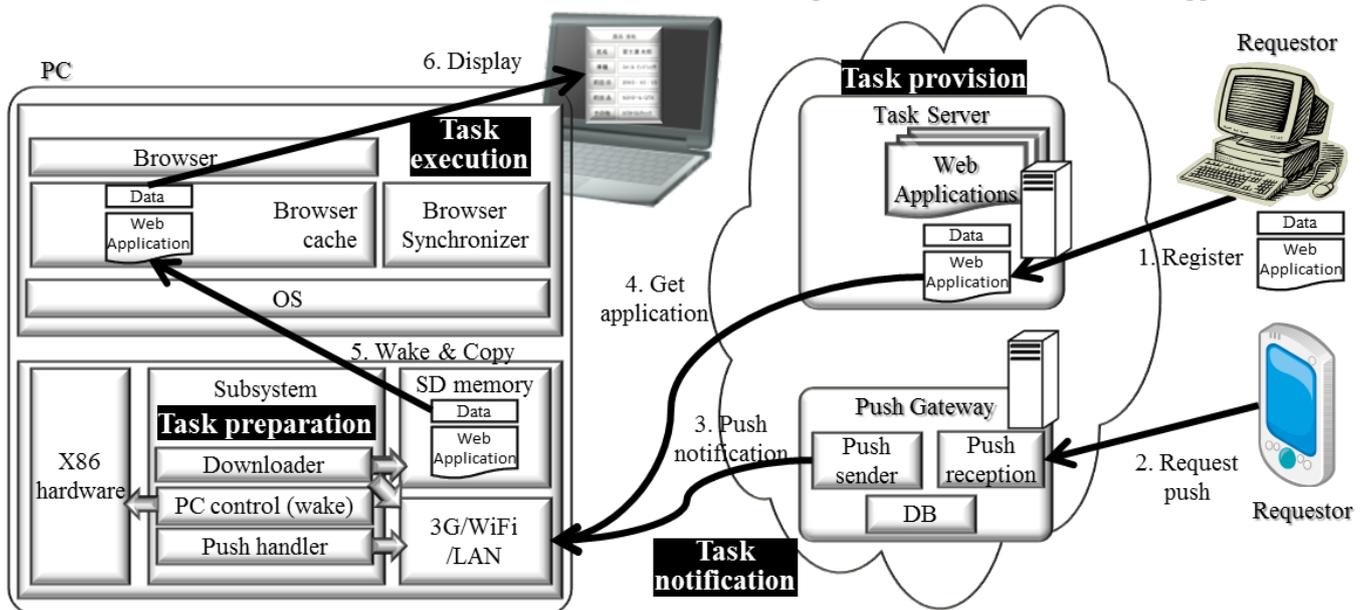
Figure 2: Prototype system.

data. We will not exclude the local application though in accordance with the momentum of the Cloud computing, we focus on an HTML5 web application here. A HTML5-based Web application can be comparable with a native application because of its higher functionality than ever before. One of the curious features of HTML5 is the application cache. The application cache provides a capability to run the Web application in the local environment, i.e., even if the network is not available, you can still use the Web application. This is good for a person or corporate that needs to work in a restricted area, where the network connection is limited for some reason, such as in a hospital. In this case, just before you get into that area, if you could receive the latest data from the Cloud service, you can use the application or data like a local application later on.

Figure 3 shows an example of an HTML5 application. To use a locally accessible Web application by HTML5, a list of resources need to be specified below the line of "CACHE:" in a manifest file. By specifying it, you can use the application in a local situation.

When an application runs, basically, it first confirms the network connection at the runtime, and if there is a network connection, then it can act as conventional Web application and access to the network as usual. However, if there is no network connection, it shall access to local application cache. Therefore, you can use the application any time.
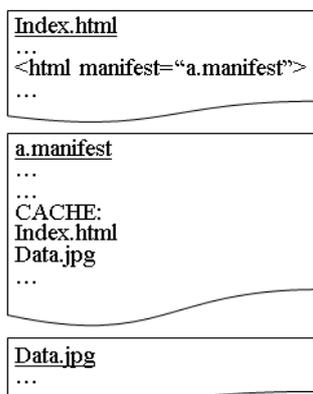
Figure 3: Web application with application cache

The detail of the four functions expressed in the architecture will be explained below.

Task provision:

Task provision provides repository, which can store application and data for download.

Task notification:

Regarding the task notification for access to computing device anytime anywhere, there are two types of notification methods: polling and pushing.

· Polling: Polling is a method by which the client device accesses the server periodically. The disadvantage of polling is power consumption. An example of power consumption by a polling method has been shown in [10]. According to the explanation, when a TCP connection to the server is established, it consumes 5–8 mA. In addition, when the data are read, it consumes 115 mA, and when the data are written, it consumes 180-200 mA. Then, a short polling consumes 0.5 mAh. If the information device polls every five minutes, it consumes 144 mAh. A typical battery for a smartphone gives 1500 mAh; therefore, about 10 % of the battery power will be consumed only by polling. Polling less frequently improves the efficiency, leading to less energy consumption. However, it loses the freshness.

· Push: Push is a way by which the server accesses the client when it wants to deliver some data. There are two types of push: true push and pseudo push.

- True push: SMS or the type of messaging in Blackberry is categorized in this type [11]. SMS is a popular method for not only communicating between users but also initiating device management such as OMA DM (Open Mobile Alliance Device Management) [12]. OMA DM has a DM Server and a DM Client for the services. The DM Server can use a notification, and it is allowed to use an SMS-like message as a trigger to cause the DM Client to initiate a connection back to the DM Server. The problem of using SMS is that it can only accommodate short size of information. For example, if you want to put a certificate in the SMS, it cannot accommodate that much information. E-mail notification is another method even though it has more than enough capability to control those computing devices.

- Pseudo push: Pseudo push notification using persistent connection has been used to control smartphones or tablets. Basically, there is no limitation on the message size. Therefore, it can resolve the shortcomings of SMS.

We decided to use Pseudo push utilizing a TCP connection to notify a task. However, it has not been established for PCs yet. Then, we accommodated necessary capability needed for the client side in the subsystem, which works independently with the PC.

Task preparation:

Task preparation is a background function of introducing the subsystem, which means that it works independently with the PC.

Figure 4 shows the flow diagram for task preparation. First, the subsystem checks the attached network device and tries to connect through a valid network device. After it establishes a connection with the push gateway, it waits the incoming push message. When a message comes, it analyzes the message whether it is valid message or not. If it is valid and the message is requested for download, then it will start downloading by accessing the given http address from the task server. If the message has a wake request, PC control function issues wake to the PC after the download is completed. It can be put additional check before issuing a wake. For example, keep the wake request until the user approaches the PC, and if it detected by a sensor, it will issue wake to the PC.
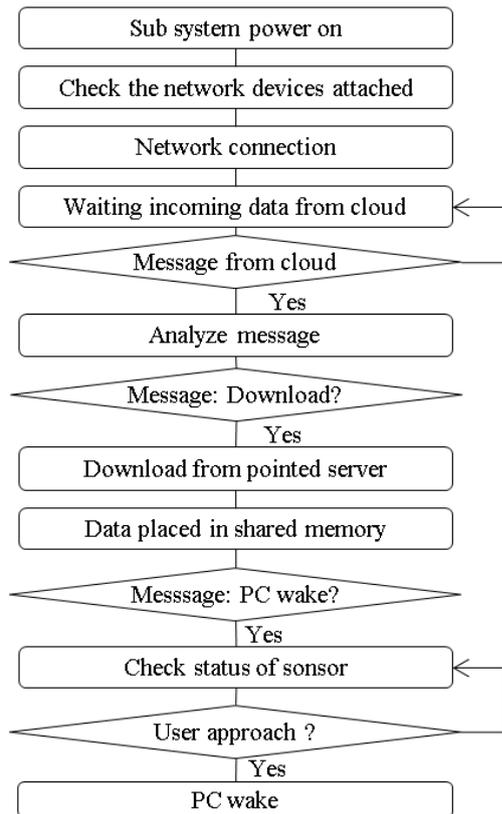
Figure 4: Flow of subsystem

Task execution:

Task execution works when PC is in on state. An added software function is the browser synchronizer. As shown in Figure 5, it checks the SD memory whether a valid application or data exist or not. If it exists, then it will reconstruct and inject the application or data to the browser cache and then initiates the browser. Depending on the request in the push message, it initiates the browser. There are two types of view on the browser. One is displaying the application icons, and the user initiates an application implicitly. Another is running the application without user interaction.
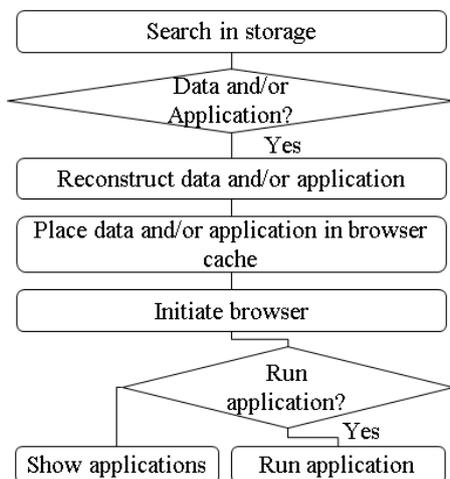
Figure 5: Flow of data collection and cache

5 SYSTEM EVALUATION

In this section, first, we show the hardware and software environments that are used for the evaluation, the result of the basic operation, and the measured power consumption.

5.1 Environment for System evaluation

Followings are the hardware and software configurations, which are used to evaluate the system.

Task Server:

· Hardware: A laptop computer is used as the task server and push gateway. It is a Fujitsu LIFEBOOK E780/A with an Intel Core i7 Processor 620M 2.66 GHz, 4 GB of main memory, and160-GByte hard drive.

· Software: A Web server is used as the task server. The public area of the server is used for placing an application and/or data. The requestor will place an application and/or data.

Push gateway:

· Hardware: One computer is used for the push gateway and task server. See above.

· Software: Linux is used as the OS. For the push reception, apache and Java Application Server (tomcat) is used. On the top of the Java Application Server, push reception is placed as a servlet. For the push sender, C-based program handles the communication with the PC. Between the two functions, those are communicating with a socket.

PC:

· Hardware: A laptop computer is used as the host system. It is a Fujitsu FMV-BIBLO NF/C80N, which had an Intel Core 2 Duo Processor P8600 2.4 GHz, 4 GB of main memory, and 80-GByte hard drive.

· Software: Three combinations are used as software environment as follows.
  - Microsoft Windows 7 as the OS and Google Chrome as the browser are used.
  - Linux as the OS and Google Chrome as the browser are used.
  - Google Chrome OS is used as the OS and the browser.

Chrome disk cache [13] is used as place to put the Web application at all above.

Subsystem:

· Hardware: A Keil MCB2388 Evaluation Board is used as the subsystem. This has an ARM7 family-based processor. It also has a USB interface supporting USB devices and USB OTG/Host, SD/MMC memory card interface, and 10/100 Ethernet interface. The subsystem is connected to the host system through the USB device interface and to the network device through the USB host interface. In terms of electrical hardware specifications, the supply voltage of the board is 5.0 V and the typical current is 65 mA with the maximum current being 120 [mA] on the basis of the specification sheet.

· Software: The board comes with RTX Real-Time Operating System, which allows programs that simultaneously perform multiple functions to be created.

MDK-ARM Microcontroller Development Kit is a software development environment that has a TCP Networking Suite, USB device, USB host stacks, and other programming libraries.

Network devices:
- ・ Hardware: Three network devices are used as follows.
  - A 3G communication device FOMA A2502 HIGH-SPEED is used as the network device. In terms of hardware specifications, the maximum downlink data rate is 7.2 Mbps and the maximum uplink data rate is 384 Kbps. It is bus powered through the USB port. The voltage is 5.0 V. The maximum current is 650 mA, the average current is 440.6 mA with the maximum standby current being 60 mA. The average standby current is 54.7 mA.
  - A LAN chip on the subsystem is used.
  - A WLAN device LANTRONIX WiPort, which supports IEEE802.11b/g, is used by connecting to the above LAN interface.

5.2 Confirmation of basic action

We evaluated the system using the Google Chrome browser on Windows and Linux and Google Chrome OS. As a task, a Web application-based photo viewer and pictures are used. In the communication device, 3G, LAN, and WLAN devices are used.

Then, the following basic action is confirmed, as depicted in, on the basis of the flow shown in Figure 2.

Figure 6: Task send and execution.

(a) User takes picture using a smartphone. (b) User sends the data to the task server by e-mail. (c) The task server makes a package of data and an HTML5 photo viewer. In addition, the push gateway receives a push request from the phone, and pushes a message to the PC. (d) The subsystem receives the push message and downloads the application and data. Then, wake command is issued to the PC. (e) When the PC wakes, the application and data are injected to the application cache of the Google Chrome browser; finally,

the web application and data are executed. At this moment, there is no network connection with host PC but subsystem. Therefore it could confirm that the application is locally executed.

Another simple basic action is confirmed as depicted in Figure 7 and as follows.
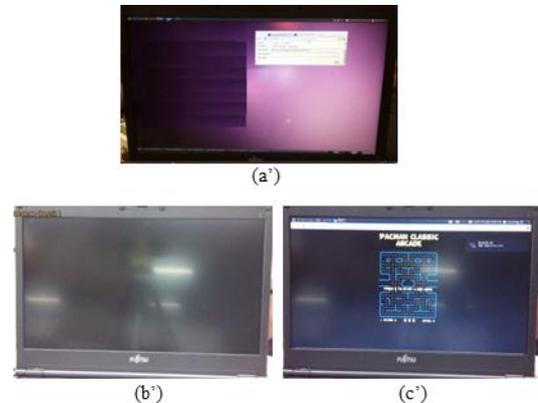
Figure 7: Task send and execution.

(a') User put a URL which shows link to a HTML5 game application at a tool on PC and the tool register the app to the task server. Then it asks a push to push gateway with a message for execute the application. (b') PC is in sleep. The subsystem receives the push message and downloads the application. (c') PC wakes. The application is unpackaged. Then it is inject to application cache of the browser. Finally a user can play.

5.3 Measurement of power consumption

We will show the results of measurements of power consumption using the 3G module. First, we examine the persistent connection. Then, we compare the conventional PC and the PC in the proposed method. In addition, we will show the line of balance, which expresses how long the system should remain in sleep mode for reduced energy consumption.

How to put the PC in sleep mode is easy; the most used way might be timeout. It can be set in the operating system. If you specify a limitation time, and no activity has taken place in that time, the PC goes to sleep without any interaction. We have used this method to put the PC in sleep.

To express the power state, we refer the ACPI (Advanced Configuration & Power Interface) [14] that defines system and device power states. Some of them are as follows: S0 is in the system working state. S3 is a state in which the processors are not executing instructions and Dynamic RAM context is maintained. S4 is a state where the DRAM context is not maintained and all devices are in off state. S5 does not save any context. D0 is a state in which a device is in the operating state. D1 is an intermediate power state whose definition varies by the device. D3 is a state in which the device is powered off. If we use the term "sleep" in this paper, then it implies that the PC is mainly in S3 but not exclude S4 or S5. If we express Network device in D1, then

it means that the network is connected but there is no communication.

5.3.1. Persistent connection.

We discuss the power consumption of a persistent connection in this subsection. We use a TCP connection to notify a task. Therefore, the power consumption of the persistent connection is a curious factor in terms of feasibility. It is expected that by using an open TCP connection without any transmitting or receiving of data, the power consumption will be low on the basis of the standard hardware capability.

First, we measured the current of the subsystem that is in D0 (On), as listed in Table 2. Somehow, it does not match with the specification described in 5.1, and the result of actual measurement is more than that of the specification. For reference, the voltage is taken as 5 V. The evaluation board used as the subsystem has an unwanted device for the evaluation, such as an LCD; therefore, it can be a small value if we use the embedded CPU alone.

Table 2: Subsystem

Subsystem status	Current [A]
D0 (On)	0.152

Then, we measured the current of the subsystem with the network device as shown in Figure 8, and the average current values are listed in Table 3.
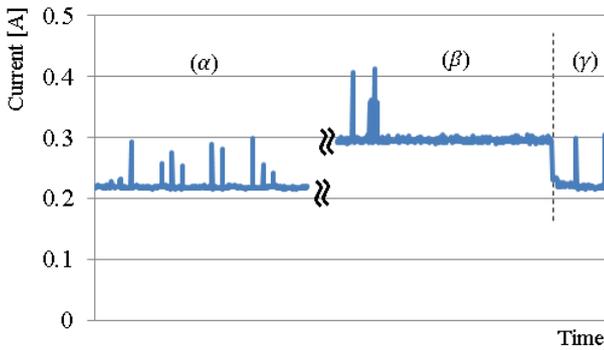
Figure 8: Measured current of subsystem with 3G.

The current changes until 60 sec and at 60 sec or later. As the device automatically transits to a low state when no explicit data are sent or received, the communication is maintained. The difference of current between the first row "plugged but not used for communication" and the third row D1 in Table 3 is trivial. Even if the network is in communication ready, it does not consume a large amount of energy. This shows that a persistent connection can be fairly efficient in terms of power consumption. Comparing with Table 2, difference between results given in Table 2 and the third row D1 in Table 3 is 0.07 A, and it is found that the current is slightly high somehow when compared with the average standby current that is described in section 5.1. For reference, the voltage is taken as 5 V. This time, we use a product of communication module for the PC, which started to sell in 2007. However if we could use average

type of communication device used in a cellular phone or M2M (Machine to Machine), the consumption will be around 1/8.

Table 3: Measured average current of subsystem with 3G

3G status	Subsystem status	Average current [A]
Plugged but not used for communication	D0	0.221…(α)
D0: Connected and idle (until 60 sec)	D0	0.296…(β)
D1: Connected and idle (60 sec or later)	D0	0.222…(γ)

5.3.2. Comparison of the conventional PC and proposed one.

We measured and compared the conventional PC and the PC used in the proposed method. Table 4 shows measurement result of average power consumption using the conventional PC with 3G. While the system is in S3, the absence of a network can considerably reduce the power consumption, which is well known.

Table 4: Without a subsystem: PC with 3G

PC status	3G status	Power [W]
S0: Idle	D0	26.45
S0: Http communication	D0	26.71
S3: Sleep w/o network	D3	2.030

Figure 9 shows the measurement results, and **Table 5** lists the average power using the proposed PC with subsystem and 3G.
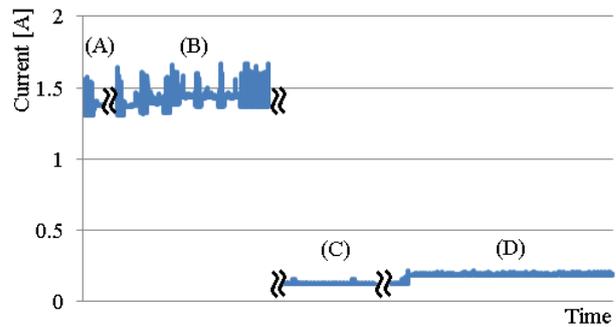
Figure 9: Measured current of proposed system.

Table 5: With a subsystem: PC with a subsystem and 3G

PC Status	Subsystem status	3G status	Power [W]
S0: Idle	D0	D0	25.8…(A)
S0: Http communication	D0	D0	26.8…(B)
S3: Sleep w/ network	D0: Idle	D0	2.41…(C)
	D0: Http receive		3.55…(D)

The power consumption in S0 is almost the same. However, in S3, power consumption is slightly increased because of the communication capability. Figure 10 shows a border line derived using equation (10) in Appendix when T = 24 and based on the data given in Table 4 and Table 5. It shows whether the proposed system can provide the merit regarding the power consumption. The horizontal axis in the figure indicates the utilization time of the conventional PC, and vertical axis shows the bifurcation point, which indicates whether the proposed system is paid off. The left hand side of dotted line shows the invalid area. If it could increase the sleep time by reducing the time to go to sleep, then it can be reached to the line. For example, suppose you use 12 hours per day currently, reducing 0.85 hour is the point matches to the current power consumption. If we could use a power-efficient subsystem and network device as described in the previous subsection, the borderline will move to much easier portion.

This is strict and may not be so adequate to evaluate the merit because a computing device that is not networked cannot receive any crucial and valid service. Therefore, it might be accepted without any concern by the consumer even when additional power consumption is involved.
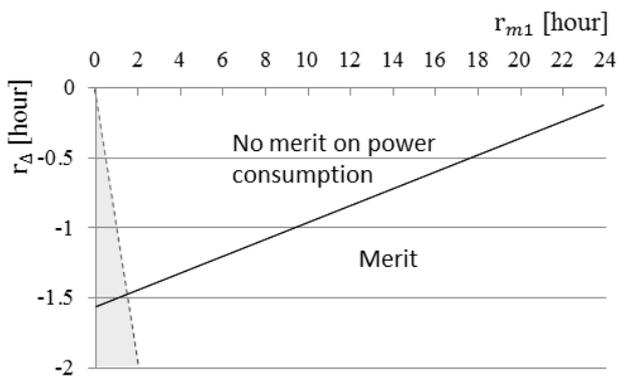
Figure 10: Line of balance of the proposed method.

6 CONCLUSION

We proposed an architecture where the system allows sending and executing of a task at any time in a PC. In our prototype implementation, we use a subsystem to keep network for dealing the pseudo push. In addition we implemented a push gateway in the Cloud and a push handler and a browser synchronizer in PC.

The system could provide value that was not provided in current PC that:

- When a user wants to use the PC, the preparation is already completed.
- Task is presented to the user without user interaction.
- The power consumption can be reduced because during preparation, the PC can be in sleep state.

In addition, we evaluated in real field using Google Chrome as the browser, Linux and Windows as the OS, and Google Chrome OS as the browser and the OS. For the network devices, we use LAN, WLAN, and 3G as the network devices.

Then, we confirmed that the system can send and execute tasks without user interaction even the PC is in sleep and pseudo push works efficiently in the power consumption point of view. From the measurements data, only small amount of additional power 0.001 [A] was required to deploy connection for push.

With regard to the energy, as the subsystem is always powered, the lower the power consumption of the subsystem, more efficient is the system. The total power consumption per day would be comparable with conventional PC when if we could reduce the use of proposed PC 0.85 hour (suppose you use 12 hours per day currently). However, the subsystem used in this study was not so efficient because of some glue logic. If we could use devices that eliminated wasteful logic or a new one, then the power consumption will be decreased.

In this study, we only focus on the PC. However, the same architecture can be used for embedded systems without any modification. In addition, the push gateway is designed to be information device agnostic. Therefore it can provide the push service to smart phone, smart tablet and other device as well.

As for the future research, remaining challenge might be how the system identifies the necessary task and the timing to push based on the activity of the users.

REFERENCES

[1] PC ENERGY REPORT 2009 UNITED STATES, UNITED KINGDOM, GERMANY, http://www.1e.com/energycampaign/downloads/PC_EnergyReport2009-US.pdf

[2] HTML5, http://dev.w3.org/html5/spec/Overview.html.

[3] Chromium an open-source browser project, sites.google.com/a/chromium.org/dev/Home.

[4] Chromium OS, http://www.chromium.org/chromium-os

[5] Yuvraj Agarwal, Steve Hodges, Ranveer Chandra, James Scott, Paramvir Bahl, and Rajesh Gupta, Somniloquy: Augmenting Network Interfaces to Reduce PC Energy.

[6] Apple Wake on demand, http://support.apple.com/kb/HT3774

[7] Introducing Intel Anti-Theft Technology version 3.0, http://antitheft.intel.com/Anti-Theft-30.aspx

[8] Android Cloud to Device Messaging Framework, code.google.com/android/c2dm/index.html.

[9] Kazuaki Nimura, Hidenobu Ito, Yousuke Nakamura, Zhaogong Guo, Kouich Yasaki, Takafumi Miura, An Implementation of Transparent Network Subsystem for PC Manageability, RTCSA 2011, 17th IEEE International Conference on Embedded and Real-Time Computing Systems and Applications, 2011.

[10] Debajit Ghosh, Building Push Applications for Android, Google I/O 2010.

[11] Simon Judge, What is Push?, http://www.mobilephonedevelopment.com/archives/832

[12] OMA Device Management Notification Initiated Session, www.openmobilealliance.com/Technical/release_progra

m/docs/DM/V1_3-20101207-C/OMA-TS-DM_Notification-V1_3-20101207-C.pdf.

[13] Chromium disk cache, www.chromium.org/developers/design-documents/network-stack/disk-cache.

[14] Advanced Configuration & Power Interface, http://www.acpi.info/spec.htm

APPENDIX: FORMULATION OF POWER CONSUMPTION

In this section, we formulate the amount of electric power consumed by a conventional PC and the proposed PC, and then derive the equation of line of balance. As described in Figure 11, the proposed PC has an additional component called a "subsystem," which is not present in the conventional PC. "Host system" is a main part of the PC. The subsystem affects the host system and the "network device." The network device is used for the communication. Each part consumes electric power.
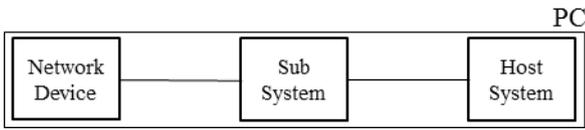
Figure 11: Proposed PC architecture.

Then, we formulate the power consumption as follows.

$$x_{on_{ave}} = \frac{\sum_{t=1}^{n_1}\big(v_c \cdot i_c(t) + v_h \cdot i_h(t)\big)}{n_1}$$

(1)

Here, $x_{on_{ave}}$ denotes the average electric power of a powered-on PC without a subsystem. n_1 denotes the number of measured points. t means the time at which the measurement is made. v_c denotes the voltage of the communication device. i_c denotes the measured current of the communication device at time t. v_h denotes the voltage of the host PC. i_h denotes the measured current of host PC at time t.

$$x_{slp_{ave}} = \frac{\sum_{t=1}^{n_2}\big(v_{h_{slp}}(t) \cdot i_{h_{slp}}(t)\big)}{n_2}$$

(2)

Where, $x_{slp_{ave}}$ denotes the average electric power of the sleeping PC without the subsystem. n_1 denotes the number of measured points. $v_{h_{slp}}$ denotes the voltage of the host PC in sleep mode. $i_{h_{slp}}$ denotes the measured current of host PC in sleep mode at time t.

$$r_{m1} + s_{m1} = \text{T [hours]}$$

(3)

Where, r_{m1} denotes the time of on state of the conventional PC in one day. s_{m1} denotes the time of sleep state in the one day of conventional PC. T denotes the total amount of time in a certain period, such as a day.

$$X_{conventional} = x_{on_{ave}} \times r_{m1} + x_{slp_{ave}} \times s_{m1}$$

(4)

$X_{conventional}$ denotes the total electric energy of the conventional PC at certain time T.

$$W_{on_{ave}} = \frac{\sum_{t=1}^{n_3}\big(v_c \cdot i_c(t) + v_s \cdot i_s(t) + v_h \cdot i_h(t)\big)}{n_3}$$

(5)

Where, $W_{on_{ave}}$ denotes the average electric power of a powered-on PC with a subsystem. v_s denotes the voltage of the subsystem. i_s denotes the measured current of subsystem at time t.

$$W_{net_{ave}} = \frac{\sum_{t=1}^{n_4}\big(v_c \cdot i_c(t) + v_s \cdot i_s(t)\big)}{n_4}$$

(6)

Where, $W_{net_{ave}}$ denotes the average electric power of the sleeping PC with a subsystem that is connected to the network.

$$r_{m2} + s_{m2} = \text{T [hours]}$$

(7)

Here, r_{m2} denotes the time of on state of the proposed PC in one day. s_{m2} denotes the time of sleep state of one day of proposed PC.

$$W_{proposed} = W_{on_{ave}} \times r_{m2} + W_{net_{ave}} \times s_{m2}$$

(8)

Where, $W_{proposed}$ denotes the total electric energy of the proposed PC in certain time T.

Obviously, adding the networked subsystem increases power consumption. Therefore, it is important to understand when it turns out to have merit. If we can put the PC in sleep mode for more time, then it can be paid.

$$r_\Delta = r_{m1} - r_{m2} = s_{m2} - s_{m1}$$

(9)

r_Δ denotes the time difference between how much spend the PC on at conventional PC and proposed method.

Consequently, the following formula can identify what time should the proposed system be kept in sleep mode than the conventional one if we wish to see the merit of the propose method.

$$r_\Delta \leq \{x_{on_{ave}} \times r_{m1} + x_{slp_{ave}} \times (\text{T} - r_{m1}) - W_{on_{ave}} \times r_{m1} - W_{net_{ave}}(\text{T} - r_{m1})\} / \{W_{on_{ave}} - W_{net_{ave}}\}$$

(10)

Session 6:
Social Studies
(Chair: Tomoya Kitani)

Practice of making friend by Japanese FaceBook and the problems

Kouji Yoshida†, Isao Miyazi††, Kunihiro Yamada †††, Satoru Fujii ††††

†Department of Information Science, Shonan Institute of Technology, Japan
††Faculty of Informatics, Okayama University of Science, Japan
†††Professional Graduate School of Embedded Technology, Tokai University, Japan
††††Matsue National College of Technology
yoshidak@info.shonan-it.ac.jp,miyaji@mis.ous.ac.jp, yamadaku@tokai.ac.jp, fujii@matsue-ct.ac.jp

Abstract— People are using twitter or blogs, but FaceBook (FB) has recently gathered attention as a social networking service (SNS) in Japan. It has many functions, as blogs do. Furthermore, photographs, information from YouTube, and web pages can be added easily. Email and chat services are available. The wall metaphor enables a user to share information with people who have become friends. Additionally, users can paste links on the wall merely by sharing. A user can communicate immediately with unknown people either in Japan or in other countries using this function. As FB users have increased worldwide recently, applications of FB have increased. However, some are spam messages, so caution is warranted.

Various problems have arisen because Japanese FB has not adjusted its systems to Japanese culture. It is surely convenient to be able to communicate easily with people all over the world, and it might broaden a user's perspectives. When used properly, it is an effective tool for various tasks.

Keywords: SNS, Facebook, blog, YouTube, Twitter

1 INTRODUCTION

FaceBook (FB) had become a topic in various ways in Japan by the end of 2010[1]. It was apparently easy to make an e-mailing list, so the author registered for membership in FB. Once registered, a user receives an email from FB, encouraging the user to be active. A user is encouraged to look for people the user knows in real life in various respects and to become friends with them.

In that way, the number of friends of a user's friends increases. However, the number of people a user can actually know remains limited. Therefore, many users start working beyond a certain geographic territory and attempt to make friends with strangers. When such a search works well, the number of friends can easily become around 50. FB becomes more interesting as the wall is enhanced with other features and information. When a user introduces photographs, YouTube movies, topics, and websites, a user can increase communications. The author currently has 1,118 friends (June 30; Fig. 1).

2 BASIC FB ITEMS

The basic points to use FB are explained below [2] [3].
Bacic FB structure is displayed at Fig.2.
(1) Basic information and a photograph must be registered as a user profile.
Register a user photograph by all means. Otherwise a person might be treated only as a newcomer. The photograph that is registered becomes a user's public face, so a user must use one that reflects a person's image. In addition, basic information is exchanged when someone becomes your friend, it is important to describe oneself well.

Fig. 1. Screenshot of a personal wall of FB.

(2) The privacy exhibition level must be set to control contents to show information depending on the viewer.
It is unsafe to show information to the general public. It is therefore better to set and limit partners to whom personal information is shown.
(3) Look for a friend
A user should find real friends where possible. However, friends a user might actually know might number only 20–30 in FB. Therefore we can use a search function and a friend introduction service offered by FB. In doing so, it might be possible to find people who are known or people who have something in common and thereby grow one's own circle of friends.
(4) Management of a newsfeed
The newsfeed, a user's own management screen, is viewed the most often by a user. It shows all information about posted information of a user and friends. Information of the newsfeed can also be customized and managed.
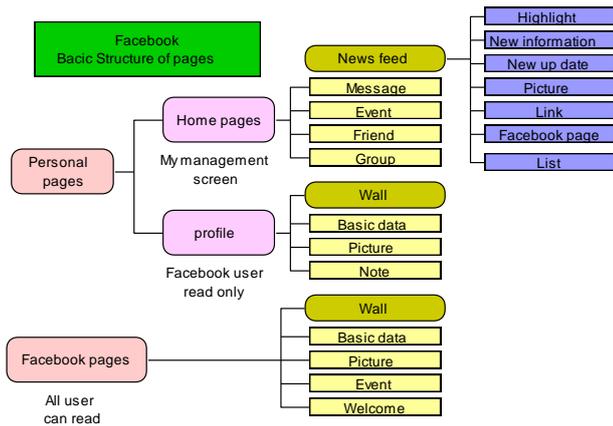
Fig. 2. Basic FB structure.

3   ANALYSIS OF PROBLEMS AND PRESENT CONDITIONS

Many people are becoming friends virtually, but without having any deep communication ("Like!", exchanged emails, notes and comments of a wall), so they can not be regarded as actual friends. Therefore, a warning comes from FB when a user tries to make friends. The warnings sometimes ask if the user and contact are real friends. If a user tries too hard to make friends, and FB suspends that friendship, a user's base points decrease precipitously. FB might even forbid a user to make friends. This system shows a considerable gap separating FB and Japanese backgrounds. FB is apparently not dealing with current conditions in Japan.

As for FB in Japan, many people from culture-related fields such as heads of media companies and small companies (especially female presidents and representatives), coaching staff, management consultants, accountants, and patent attorneys are registered. However, people from physical science fields are few. It is considered that because e-mail and access of WWW is under surveillance, especially in engineering departments of big companies, they can not use FB at their workplaces. They register themselves using personal email addresses at home, which keeps the number of such users low. World users of FB is displayed at Fig.3.

Fig. 3. World users of FB.

3.1   How to increase your FB friends

Then how does a user increase FB friends? The general method is to make friends with real friends, which can eventually number around 20–30 friends. However, how does a user increase friends thereafter? A good method is communication with FB friends, and friends of friends. Particularly, if you click the "Like" button, you are already communicating with the person. In addition, if good photographs and videos and comments are shared from your friends on your own wall, and friends who saw them give you a "Like" message, your points increase.

Then, if a user tries to make friends with people who gave a "Like" or who received a "Like" from the user, it is easy to get approval as a friend.

Here are examples of potential new friends.

(1) One who shares photographs and movies and messages of his friends, and gets many "Like" buttons clicked.

(2) One who posts photographs and movies, and gives many comments on the "Like" button clicked.

(3) One who posts a Twitter haiku and comments of a photograph daily, and gets many "Like" buttons clicked.

(4) One who posts a saying or a proverb of a calendar every day, and gets a "Like" button clicked.

3.2   What increases FB friends?

Every user must register a photograph for the initial profile. Without a photograph, a user would be treated as a newcomer at FB. Next, it is important to describe basic information well. It is at least necessary to specify an occupation and a location where the user resides. In addition, one can provide an educational background and current occupation with an acceptable level. "Self-introduction" of basic data is also an important part. In this area, size does not seem to have a limit. Therefore, it is important to complete this information efficiently. People who have good walls write "Self-introductions" well.

"My method to increase 100 FB friends in a day." by Ashikaga Nicoann school[4].
"A photograph" and "an image" have much better impact than only a message.
We can post anything such as photographs, images and animations to FB.
(1) Cute type (a pet, a child, a baby)
(2) Food type (a lunch, fruit, a cake)
(3) Nature type (a blue sky, the sea, the morning sun, setting sun)

Particularly, children and animals have been unbeatable for finding friends from the very beginning. When a user posts photographs of dogs, cats, or babies, the user is sure to get more numerous "Like" counts than could reasonably be expected.

Check your photograph on the profile again. It must be a nice, favorable one even if not necessarily taken by a professional photographer.

Fig.4 is my actual FB wall.

Fig. 4. Figure of an actual FB wall.

3.3　Problems and solutions that arise when increasing FB friends

When a user applies to make friends, a message must be attached even though it is troublesome. A point will be added if a user does so. In addition, if a week passes after a user's application and the prospective friend continues to ignore the user, then the user should go to the wall of the prospective friend (it remains in a user's messages while sending a message) and click "Cancel friend application"; so that the user can reduce the suspension of points that would otherwise occur[10].

[Suspension of friend application　Algorithm analysis]
"FaceBook pulling in customers" lecture making

[Assumption]
FB analyzes things that an individual user clicks as an "algorithm" and guesses each user's preferences. Management by inspection is impossible for suspension of applications for making friends because the number of users is about 600,000,000. Therefore, FB manages suspension work using automated programs with a constant numerical formula.

[Tendencies of people who get suspended when soliciting friends]
Non-active users (photograph unregistered or few postings) are easily suspended. Sudden suspension can result from having sent applications for making friends consecutively. However, some users do not get suspended even if they apply for 300 people a day, whereas others get suspended with only 20 applications in a day. A period of suspension becomes longer as 2 days, 3 days, 7 days, 14 days, and 30 days. Applications sent to foreign countries can be suspended rapidly. In addition, warning information is sent to receivers of friend applications recently (Nemo O'zeky reporting).

[Algorithm of making friend suspension]
FB judges suspension automatically mainly using three numerical formulae: a point addition method, a point deduction method, and change of points over time.
(a) Added points:

When clicking "Like" (+1 point), when giving a comment (+5 points), when clicking a "Like" button (+5 points), when getting a comment from someone (+10 points), and when making a remark (+3 points).

(b) Deducted points:
When making an application for making friends (-20 points), when receiving an application for making friends (-3 points). However, there are conditions that can help mitigate deductions.
* in cases of belonging to the same FB group.
* in cases of belonging to the same group.
* in cases of posting comments or clicking "Like" buttons for each other.
* in cases of exchanging messages.
　Deduction points seem to decrease greatly when people have mutually communicated in the past. Deduction points can be reduced approximately 50% when a user belongs to the same group as a prospective friend, and can be reduced 80% when a user has exchanged messages with others in the past. In other words, a user can get only 4 points deducted if the user has sent direct messages to a prospective friend before a solicitation, but might get 20 points deducted when sending a solicitation to a prospective friend with whom the user has never communicated.
　With added points (a) and deducted points (b), when a user's total negative points exceed a certain threshold in a certain period of time, a user is suspended from making friends.

(c) Changes of points by time
As time passes, additional points and deduction points change. Even if a user has 100 points now, a user might have only 10 points seven days later. Deducted points also change over time. Applications for making friends tend to be performed consecutively at one time, and deduction points are counted along with them. For this reason, people easily get suspended from making friends.

[Solutions]
As implied above, it is important to reduce deduction points counted in the same period of time: a user had best avoid trying to make friends with overseas users with whom one does not communicate in daily life, except when a user has business intentions or other particular reasons. Lack of communication raises the amount of deducted points.
　Fundamentally, FB tries to connect users realistically so a user had better avoid making or receiving applications for making friends with users to whom one can not communicate. In addition, even if a user can communicate with the prospective friend, it is better to click at least one "Like" button for the prospective friend before making a solicitation. Thereby, FB judges that a user and a prospective friend, who can give you additional points, are mutually related. However, a user can not click a "Like" button for users who are not friends yet, although it is possible.
　Therefore, it is the best to apply to make friends after clicking a "Like" button for the user commenting on an

article of a prospective friend. In this way, a user rarely gets ignored because
◆The user should be an active one
◆It is likely that the prospective friend is currently logged in
※ One receives deduction points when a solicitation is ignored or suspended.

[Summary]
A user should apply to make friends after clicking a "Like" button for a prospective friend who is commenting on an article of another user. Consequently, one gets additional points more and FB judges that a user and the prospective friend are related. In addition, a user must choose a prospective friend who is active and favorable before making a solicitation for friendship.

4 PRACTICE CONTENTS

4.1 Friend change and an actual case

・Motivation for using
The author first turned to using FB because it was troublesome to make a list of emails during authorship of a book of JAVA with four other university professors.

Date: Wed, 29 Dec 2010 19:35:07 -0800
To: Koji Yoshida <yoshidak@info.shonan-it.ac.jp>
From: FB <update+kr4m2gernnmn@ FB mail.com>
Subject: Welcome to FB - face book!
FB
Hello, Koji. Your account has been made!
With FB, exchange information and communicate with your friends easier than before.

After I registered myself, I kept the account in its initial state. FB started to send messages to encourage me to look for friends or make friends.

Date: Thu, 13 Jan 2011 16:50:37 -0800
To: Koji Yoshida <yoshidak@info.shonan-it.ac.jp>
From: FB <update+kr4m2gernnmn@ FB mail.com>
Subject: Meet your friend again at FB
Hello, Koji.
You do not seem to have been using FB very much recently. We shall introduce some acquaintants to you. Make friends with your friends and your family, your classmates, and your fellow workers and exchange information or share photographs with them.

Then I made friends with my former subordinate from the company I used to work for and with an acquaintance of a university. The number of my friends increased little by little. On January 27, I registered a list of a textbook authoring group, so eight users became my friends. When using FB in Japan "a special feature appeared in the Weekly Economist magazine[4] [5] for good looking FB pages in various ways for a week", about which I wrote a comment. Furthermore, I made friends with a user in the U.S.A. because I informed that user of the contents of an email from a foreign country and I got a "Good!" message.

I showed these contents to a Japanese woman who was interested in the information "I've got a special feature in the Weekly Economist". She became my next friend. She was the eleventh friend, actually. Then I made friends with a friend of my acquaintance in the USA on February 4; he became my twelfth friend. I tried sending a message and pasting a photograph by my cell phone on February 5. Pasting a photograph was successful, although the letter was legible only in the subject. I also found out that Android and iPhone have other applications. I also made friends with two English teachers on February 18, which means I had 14 friends then.

I started sending applications for making friends to users who were apparently positive people since around February 25. By that time, I had made 20 friends in 25 days. One of my friends' birthday was February 27. After that, I started sending a picture or a message to all of my friends on their own birthdays. Consequently, I began to realize that it is important to be a busy user at FB. I found out how to change my association status and revised it on February 28. I noticed that my friends and friends of my friends had not been able to see my page and revised it on around February 28, which came to be extremely important. I had been wondering why no one responded to the information related to my wall and I was looking for solutions. At this point, I rewrote my basic information well and changed my portrait.

I got another friend on March 1 and commented on his wall, and the number of friends increased. I made friends with Mr. Takekawa on March 3. I shared his photographs of Biei in Hokkaido with him on March 3 and earned points. I also started posting photographs I took on my trips to places such as Dubrovnik and Plitvice Lakes municipality.

I announced at my wall that I had 50 friends on March 5. "I have made friends with 50 people. It has been really nice getting to know all of you. I hope I can be helpful to you as I post information and exchange opinions with you.
The proverbs that I hit on today:
* The poor can not afford manners.
* An idle brain is the devil's workshop.
* Wise men learn by other men's mistakes.
As Weinberg says in "The Way of a Super Engineer"[7], a self is blind. Therefore, he recommends setting a mirror around you and keeping records. Then you can see yourself objectively from doing so.

"Posting situation on my wall, number of friends and topics"
I posted a note of "How to handle work 1" on March 5.
My FB friends numbered 70 on March 7.
My FB friends numbered 100 on March 9.
The Great Eastern Japan Earthquake hit on March 11.
Information of the earthquake overflowed in FB.
I started making friends with users who were from the engineering field.
My FB friends numbered 150 on March 25.
My FB friends numbered 200 on March 31.
I accelerated my pace from that point.
My FB friends numbered 300 on April 5.
My FB friends numbered 400 on April 9.
My FB friends numbered 500 on April 13.
My FB friends numbered 600 on April 18.

My FB friends numbered 674 on April 20.
Then, I got suspended for two days.
The suspension ended on April 22.
My FB friends numbered 700 on April 23.
I slowed my pace from that time.
My FB friends numbered 800 on May 5.
My FB friends numbered 882 on May 16.
I slowed my activity to write an article then.
My FB friends numbered 887 on May 18. I started my activity again.
My FB friends numbered 900 on May 22.
I began to sort my friends a little better from that time.
My FB friends numbered 1000 on June 9.
My FB friends numbered 1119 on June 30.
Fig.5 is the change at the number person of my friend.
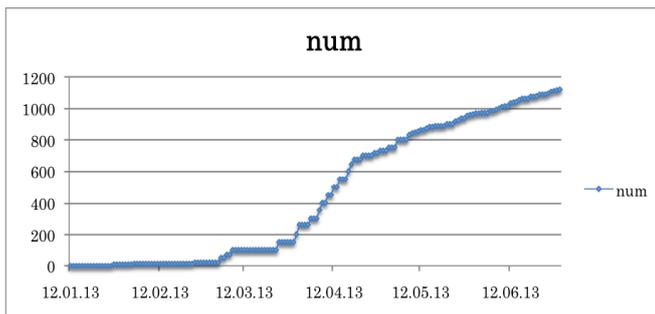
Fig. 5. Chang at the number person of my friend.

4.2 Friend history and the change

Most people have about 20–30 friends at most in real life. Therefore, it is apparently the "key" number in making friends. In other words, it can be a clue to decide whether you increase your virtual friends other than your friends in real life or not. From my experience, as one gets over 50 friends, introductions from FB also increase, and you can increase your number of friends more easily and reach 100. Once you get over 100 friends, the number of your friends can increase easily to about 700 rapidly if you do not select users and merely increase your friends. In my case, it took 40 days to increase friends from 100 to 700, which is an average of 15 friends each day.

On the other hand, a person concerned with real friends is a person who values these 20–30 friends, and friends who are not truly friendly show many people on FB who decline a solicitation to become a friend. Because a person with 200–300 friends is more than one intends to increase as real friends, by my experience, it is easy to obtain friend approval.

In addition, walls show users who want to get more friends. It is effective to make those users your target to make friends with them.

I also thought that just getting more friends is meaningless so I tried to post and share useful or interesting information. Once the number of my friends reached over 900, I started to choose users who had unique walls to make friends. Thereafter, the rate of increase declined.

5 ANALYSIS AND EVALUATION

FB is not apparently so prosperous yet in Japan. Only a certain group of people are using it. In addition, having feelings of resistance to use one's real name and a real photograph is one reason that users are not increasing in Japan very much. Furthermore, Japanese FB staff are few, and it seems that they are not dealing with Japanese culture. The user ratio in Japan remains only 1.4% of the total population (Fig. 3).

5.1 Analysis of problems

FB users are considerably fewer than those of other social network services (SNSs) in Japan, such as mixi (about 20,000,000), GREE (about 20,000,000) and Mobage town (about 18,000,000). There are about 1,800,000 FB users in Japan [8].

(1) Not having Japanese specifications.
As basic information, FB asks the location of your residence, hometown, sex, the date of birth, and blood type. However, the next question is "Is your love interest female or male?" Since you have already declared your sex, it sounds like a meaningless question. Perhaps this question is meant to show consideration for gay people.
(2) Using real names and specifications for reality
FB has a basic rule of using real names and of associating with friends in the real world. Therefore, if you have insufficient friends (over 100), you might be unable to enjoy using it. Many people in their 30s took entrance examinations for universities before universities started missing their quotas because of declining birthrates. For that reason, many of them graduated from universities of low rank irrespective of their intelligence. Furthermore, FB asks a user the names of a high school and university only, not the junior high or elementary school. It implies that they assume users have a high educational background.
(3) Giant of SNS to move ahead
Like Instagram, which allows joint ownership of photographs, many Japanese do not want to just take their relations of Twitter with FB if they can just port the human relations that have already been made on Twitter. We use Twitter in an anonymous account. In contrast, FB seems to build and maintain slow connections and heavy connections. As for friends, there are apparently many people who do not give much thought to the distinction.

The real-name principle of FB blocks its progress as a social network.

5.2 Evaluation

Why are FB users not reproducing quickly in Japan? [9].

Regarding its scale, FB is not even one of the largest SNSs in Japan. The number of FB users is greater than 500,000,000 users, and most are foreigners. To talk to them, one needs a certain level of language skill. However, judging from the whole FB general meeting numbers, there are few Japanese FB users. FB is regarded as "a tool for people who have friends to talk to in a foreign country", and

it is one factor making Japanese users keep their distance from FB.

The application function is too complicated, almost incomprehensible

As a characteristic of FB, various application functions are available, and it is superior in the sense of compatibility with joint ownership of an image and animations and other web services. However, "multifunctionality" is often inversely proportional to "plainness" or "simplicity", and as for FB, multifunctionality constitutes an important blockage for new users' entry. It is necessary to help users understand and use the contents of the multiple functions, so there must be a "plain explanation" for users. However, no such consideration exists in the Japanese edition of FB.

The reasons why FB, which should be proud of having the world's greatest number of users, is forced to have difficulties in Japan can be summarized in the following three points.

(a) Language barrier
(b) The system has too many functions
(c) Lacking structure of a support system that is necessary to guide beginners and encourage them to be active users.

6 CONCLUSIONS

One can be connected with other Japanese users who are strangers, or potential friends. It is also possible to make friends with overseas users. It has not been easy to make friends with people in other countries, but FB has enabled Japanese people to communicate with people worldwide[11][12][13].

However, if one becomes too absorbed in this world, one might spend too much time in the virtual world rather than in the real world, which can be called "FB syndrome". After all, it is important to manipulate it as a tool, and not to be manipulated by it.

REFERENCES

[1] The public time of a Japanese edition of FB
http://www.itmedia.co.jp/news/articles/0805/19/news046.html

[2] Maka Inoue,Shinichi Satou,"Complete understanding facebook with this book", Nikkei B P (2011/3)

[3] Tomoyuki Negishi,"A facebook maneuver", Ascii new book (2011/2)

[4] Ashikaga Nicoann school
http://www.facebook.com/ashi.nico.gaku

[5] "Weekly Economist magazine 2/1",Mainiti newspaper(2011/ 2)

[6] "Weekly Diamond magazine 1/29",Diamond(2011/1)

[7] G ・ M ・ Weinberg , "The load of super engineer," Kouritsu books(1991/10)

[8] Is FB in Japan popular?
http://dragoncry.blog34.fc2.com/blog-entry-191.html

[9] Problems of the FB which does not readily spread among Japan
http://www.tanautsu.net/blog/archives/weblog-entry-79.html

[10] "A face book pulling in customers" lecture

【Suspension of friend application Algorithm analysis】

http://www.facebook.com/notes.php?subj=195496437142705#!/note.php?note_id=187593784608144

[11] "Social Media strategy" , DIAMOND Harvard Business Review(2011/4)

[12] Albert-Laszlo Barabasi,"LINKED: The New Science of Networks", NHK Books (2002/12)

[13] Nicholas A.Christakis and James H. Fowler,"CONNECTED", Koudansya (2010/7)

To educate a Knowledge Resourceful Staff-CIO

Kazuya Matsudaira [*]

[*] System Frontier Corporation
mtdr@pji.co.jp

Abstract –Japanese modern organization must have a CIO in charge for IT as the vital staff [1]. The paper reports a secret of comprehensive survival guide for such many organizations as who are surviving for long years in view of KRM (Knowledge Resource Management). Staff-CIO is the vital human asset to keep strong management of enterprises. Specifically, CEO supported by a Knowledge Resourceful Staff-CIO utilizes adequate information or knowledge including a kind of management know-how for their competitive decisions making.

However, even now, most management can't get enough information through use of IT. The key to this solution is to have the right Staff-CIO with adequate career and his/her school education backgrounds. The Staff-CIO requires variety of knowledge fields. An academic degree will be distinctively advantageous. Specifically, Staff-CIO may have to supplement this with knowledge of two cultures–the humanities (literature, history, culture, and liberal arts, so on) and the sciences (mathematics, physics, chemistry, engineering, so on). Knowledge sensitivity is necessary to estimate the social and human consequences of business decisions. Scientific knowledge will be required to distinguish between good and bad experts.

A lot of advises for the CEO is replete with disagreements between experts [2]. To educate and get help by a complete Staff-CIO, CEO must understand, as history teaches us that two more areas, in addition to the Humanities and Sciences, are to be added. One is Human Factors. The other is Art of War.

Keywords: Knowledge Resource Management, CEO(Chief Executive Officer), Staff-CIO(Chief Information Officer), Human Factors, Art of War, Humanities

1 Introduction

Henry Ford said "There is one rule for industrialists and that is: Make the best quality good possible at the lowest cost possible, paying the highest wages possible [3].
Most managements know it, but not easy to practice. Modern Management knows they need Staff-CIO to solve problems. To describe the duties and responsibilities of Staff-CIO, some analogy of driving car is useful. To be no Car-crash or no mistakes of driving by cab-drivers in some long years, those cab-organizations must have some intangible assets like philosophy, mission, or vision that helps manage company to alive in long life.

A Japanese typical long-life company who were founded in old age has possessed value-driven knowledge resources as a strategic enabler. At such a company, we can find the catalyst or evangelist to keep communicating important discourses in the organization. They are said the chief staff is a "Bantou(

)". Banto is the Staff whom the owner of the company relies upon. Banto has always ABACUS (old IT tool) with him and keeps the shop profitable and solves all the problems. KONGOUGUMI, the oldest company founded in Osaka, Japan as constructor of Shitenoji(　　　) temple in 593 AD. The company has 1408 years-life and it is still vivid in Japan. The company currently continues construction business and some jobs of maintenance of temples and shrines. The company experienced many of difficult periods since founded in Osaka. At that time, they were able to access to Kyoto as the center of construction business very easily. But, Tokyo has become center of business about 140 years ago.

At World War , their offices were destroyed by fire, and during recession of 1980 decade, the company faced money problem. Their business is architect-jobs of classical temples/shrines. CEO must know how to design and construct such a huge temples and shrines. Most of their employees are craftsmen and carpenters. They listen to CEO's narratives of orders in severe situations. They can believe occurring problem faced with them is to be able to get solved. Actually, the CEO's narratives give a power and brave for the employee to challenge to correct their difficult business. The story is told eye to eye, mind to mind, skin to skin, and heart to heart among employees. The CEO has in this way protected the secret of special carpenter skills, then long life of the company is assured. In the long years, he honestly follows same behavior as previous CEO, his predecessor. He repeats to talk to employees on organizational discourse, or narrative when he shares dialog time with his employee on business concepts and philosophy in a manner of a story-telling by his mouth.

This way is deemed as one of unique management style in Japan. Practice of intangible knowledge by collaborative communication keeps company sustainable for long time. The discourse is a vital knowledge to enable employee to work [4]. The organizational discourse that has kept in more than one thousand years is a vital resource to keep KONGOUGUMI. The Staff-CIO who is "number 2" as a skillful carpenter is always helpful for CEO who is most skilled carpenter in the organization. Since the very old era, Staff-CIO has background of 2 circles arena. (Fig1). His knowledge is mainly consists of Sciences like architect and construction engineering. But he has to have broader humanities. His mind has three basics that are Ethos, Logos, and Pathos. He never forgets that he, as a Staff-CIO speaks with CEO and support CEO in any serious occasion. We should learn from the history.

Old Japanese proverb says "　　　　　　　　　".
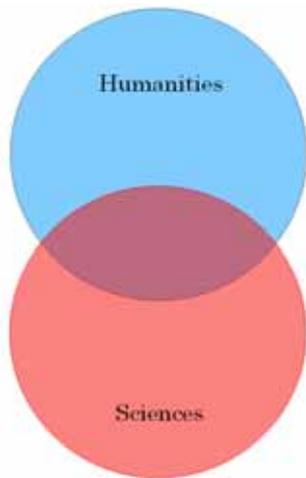
Fig1 2 Circles Model

2   A GREAT STAFF IN NARA(   ) ERA

  Long years ago, Kibinomakibi(693   775) was the famous staff of Female Emperor Koken(   ).
He went China with Abenonakamaro(   )   by the mission to Tang China and monk Genbo(   ) when he was 22 years old after educated at national institute of Daigakuryou (   ) and got more education in China for 17 years. There, he was educated in variety of fields. In Humanities arena, he studied mainly, Analects of Confucius, or Lunyu(   ), History of China, Music, the art of embroidery, Chinese Poetry. IGO(   )-game was imported from China to Japan by him.

  In Sciences arena, he studied Astrology, Mantic Art and Physics and Chemistry, Mathematics so on. In the military area, he studied the Art of War and other 6 military classics. He also got copy of the Sun Tzu, Wuzi among China's Seven Military Classic with the engineering method of Castle Architect and Construction , so on.

Pic.1 Kibinomakibi plays IGO

On his second trip to China, he came back with the famous blinded Buddhist Ganjin (   ) who taught as a Chinese to Japanese new concepts of how to apply Buddhism to governing Nation in Japan. Through these experiences, he got also a lot of religious knowledge. As a military science, Sun Tzu (   ) or Wuzi (   ) were brought back to Japan. China's seven Military classic were introduced by him, so he was the Japanese originator of military Art of War. He was the real Staff for the Emperor. He utilized the Art of War to the real war of Eminooshikatu (   ). To him, the art of war is a kind of "theories-in-use". When he discovered serious problems through the interpretation of internal and external environments according to their knowledge, he is easy to solve the shared problems among the cabinet of Government. His strategy was revealed by manner of dialogue with Emperor and those are related to voices of people [5]. It seems succeeded management is easily communicated to other staff. Serious conflicts were never happened when he supported the Emperor. Sometimes, some chief politician might complain of the rules which he set up. DENPOUKIKIGAKI(   ) is a small manual to teach the quality of politics.   Kibinomakibi is not only a Confucian Scholar, but also politician who knows politics both in peace era and bloody battle and war era. He deemed himself as Staff-CIO(   ) when he wrote the letter of resignation from the Ministry to the Emperor. Gaikotunohyou: is a letter that he wrote in AD 770 year to ask permission of resignation. He was the honest Staff-CIO who executed his responsibilities. He was also inventor of KATAKANA MOJI [7]. Benjamin Franklin said; " Remember not only to say the right thing in the right place, but for more difficult still, to leave unsaid the wrong thing at the tempting moment."

3   TYPICAL EDUCATION BACKGROUND OF STAFF-CIO

The 3 circles model is to be widely accepted to explain right staff's knowledge background in Japan. (Fig.2)
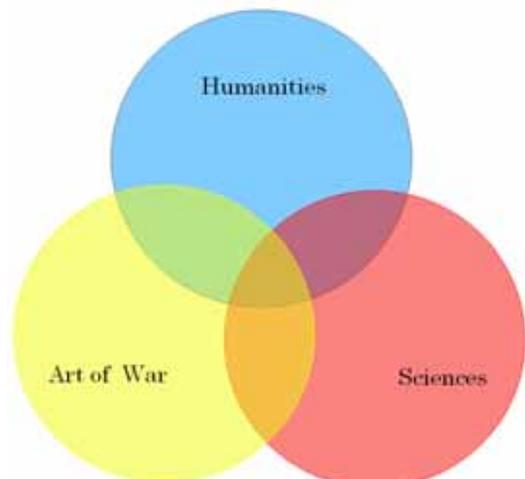
Fig.2   3 Circles Model

Kibinomakibi belongs to this model. The 3 circles model is appearing and it has demonstrated to be the good combination of education for the Staff [5]. In 3 circles model, as same as the two circles model, the first circle is named Humanities circle. The second one is named Sciences. The third one is added as Art of War. The Japanese Military Academy since built in Meiji era deemed this 3 circle model was enough to make a good officer. The Naval academy had paid more attention to the Education of Sciences than Humanities because ships were their necessary tools. The Naval Staff College was an escalator system to make a higher level of staff. The Japanese Military Staff College was an entrance gate to promote the graduates to the General as a commander course. At Shyowa era, we had failed to train the right Staff-CIO at the Army. The history told us Japan knew the defeat of intelligence in 1945. The cause of the defeat of war existed in the contents of staff education and training programs. To correct it, one more circle which is Human Factors circle should be added. We may call it 4 circles model. (Fig3)
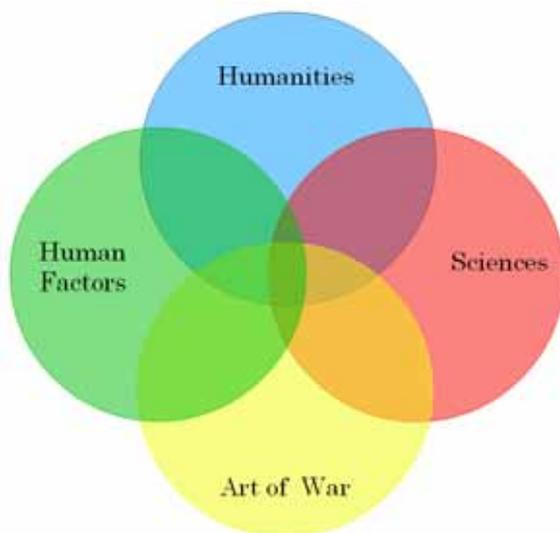
Fig.3   4 Circles Model

The Circle of Human factors is very tightly related to understand human-being of both physical and psychological side. With the background of human factors, the Staff-CIO may give straight suggestions or frank opinions to CEO. He might correct some illegal decisions/actions before seeing actual failures. In 4 circles model, there are one spot which belongs to all 4 circles. Those people who belong to this spot believe that the Oneness feeling is most important relationship with his or her CEO at the organization hierarchy. Without this team-work, the company might be destroyed by some accidents like monetary problem, or a lack of corporate governance in the severe business situations. CEO may get consultation on his health if the Staff-CIO has some skill of medicine. Of course, to become the typical staff-CIO, one may start with one circle. Gradually, each of 3 circles has been added to the candidate of Staff-CIO. Although young Japanese entrepreneur prefers to employ a established Staff-CIO with 4 circle model to build a sustainable organization, there should be a lot of difficulty because of lack of personnel resource. It is wiser that the company would educate after

finding the person with potentiality to become a Staff-CIO of 4 circle model person. Knowledge management may be failed in too soon improvement of organizational performance under the intra-organizational competition. Special anti-Staff-CIO actions may be happened. If this is caused by not sharing Knowledge, the next sound decision is to change current Staff-CIO.

3   OHMURA MASUJIRO(   )   THE IDEAL   STAFF–CIO

Masujiro Omura is the ideal Staff-CIO in the history of Japan. He is a staff from 4 circles model. He was born as a son of medical doctor, and he is to become to a medical doctor since he is a boy. He went to school of Dutch medical Doctor,Yuusai Umeda (   ).   Then, he went to the school KANGIEN of Hirose Tanso (   ) who was the famous scholar of Confucius. After that, he went again to Umeda, then went Ogata Koan academy school (   ) in Osaka. There, he learnt The Art of War, Dutch Medical and Sciences and Engineering. After he started business of Medical doctor to succeed his fathers' job, but soon he failed it. Then, he went to Uwajima Han (   ) to help modernization of War preparation. He made a actual war-ship, canons, and taught the military strategy and tactics. After he did it, he went to Edo (   ) , there opened small school of art of war, named Kyukyodo (   ). While he was in Edo, he went to Yokohama every day to study English and modern sciences and medical methods from Doctor Hepburn and his wife Clara.   He went back to homestead, Tyosyu (   ) to become the Staff-CIO for his country. He educated young officers and soldiers by his written manual in Japanese of Tactics. Tyosyu battled with Tokugawa Bakufu (   ) and by his plan of operation, Tyosyu won the first war against Bakufu.
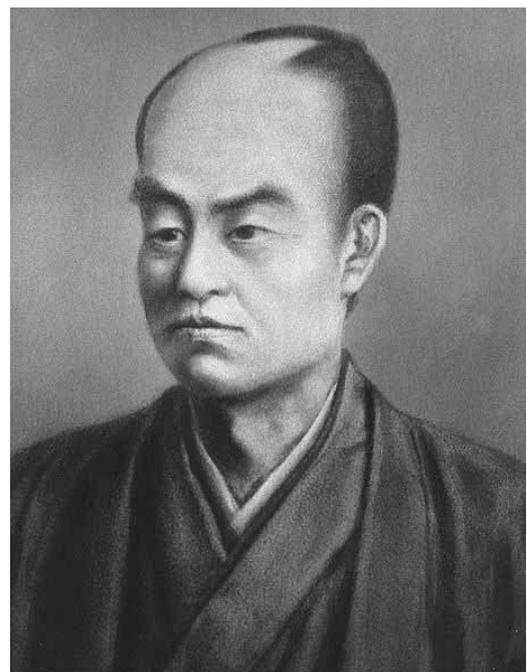
Pic.2   Omura Masujiro:

At new era Meiji, he became officially a chief of Staff-CIO(   ) to the Meiji emperor and made significant performances. But, he was killed by terrorist when he was 45 years old. Omura Masujiro had from both theoretical and practical sides, the ideal background as a Staff-CIO. He had a full spectrum of knowledge of 4 circles [6]. At Meiji era, Information and Knowledge had become vital military sources to keep Japan competitive against foreign countries specifically like Russia, Europe and USA.

Educating and empowering Staff-CIO had become one of the main themes of modern business management. They adopt new management style but continue to believe the best way is stick to old their culture. Most of them are rich company that keeps enough cash-flow. Some of them are very visionary company. Quite a few Top managements still understand importance of Knowledge/Information and fails to make employees use those Knowledge resources. We see their ancients had formulated by their experiences like below.

Wisdom=Knowledge+Experiences.

Knowledge=Information+Analysis.

Information = Data + Processing [4].

The classical principle says that there are two types of executives: line and staff. Line executives are those who supervise activities that contribute directly to the profit of the enterprise. While staff executives are those who contribute indirectly by providing services or information. The existence of Staff department makes it very difficult to keep spans of control short, and it makes it impossible to have a single line of authority from the CEO. The difficulty may sometimes exist because of higher management's neglect of definition of jobs carefully. The Staff-CIO is to be a chief among general staff,

And he is a man who has authority over the staff departments in the enterprise.

4 CONCLUSION

To pick up one more person from long history of Japanese Staff-CIO, Myoan Eisai (1141-1215) who is the originator of Rinzai-Zen (   ) in Kamakura era is ranked as No.2 Staff-CIO. He has also 4 circles background of education. He brought green tee ( ) from China and recommended it to Kamakura Syogun "Sanetomo (   )" to recover his health with the book titled Kissayojyoki (
). Eisai had deep knowledge of medical treatments and medicines.

Rinzai Zen Buddhist priest generally trained himself by training system of Rinzai-Zen. Those trainee were to learn using a thousand set of KOAN(   ) which was the disciplined example book of solutions for difficult questions about how to live. The teacher was easy to examine whether the student had understood Rinzai's concepts/philosophy that was built in Rinzai-Zen.

Pic.3 Myoan Eisai:

After Eisai died, successors of Rinzai-Zen, had been active in staffing Samurai-Syogun for about 500 years. Eisai himself was the Staff-CIO for three generations of Kamakura Shogun. After analysis of Eisai background and career, we may deem him as an original staff from 4 circles model [7]. Since Eisai, medical Doctors like Dosan Manase (   ), Zenso (   ), Tenkai (   ), all co-acted both Syogun's medical staff and information staff as a priest staff (   ). After Edo era, Tokugawa Bakufu (
) has limited free activities by Temples and Shrines. Eventually, very limited educations were only deployed at Zen-Temple. Even Ashikaga Gakko (   ) lost power of producing so called teacher-Buddhist as Human resources(   ) because of decrease of Tokugawa's support. End of Tokugawa Bakufu era, some medical doctors who got study from Netherland on modern medical and sciences. Those doctors learnt the modern art of war that was newly created by Napoleon. For example, Sugita Genpaku (
      ), Ohtori Keisuke (   ), Furuya Sakuzaemon (   ) ,all medical doctors had tried to become Staff-CIO to realize the new age at end of Tokugawa Bakufu. In Meiji era, military staff school was set up, but it's objective is mainly to make Generals/Admirals officers not for staff officers. It was unbalanced education system in terms of breeding the Staff-CIO. Both in modern governmental organization and business organization, we need the right Staff-CIO. We must notice they need broader background and deep disciplines.

To most of them, knowledge about human-being was lucked. Education on Human factors was never done. Specifically medical and psychological aspects and the mental side of human-being were to be trained while students are young and fresh. Scientific analysis of human mind is necessary factor to function company organization as same as dynamic mechanism of human body.

After World War   , they believe craftsmanship is in decline due to the general apathy fount in today's society.

Staff-CIO is a kind of craftsman at the company. Japan as a country needs the right Staff-CIO. Craftsmanship of the Staff-CIO is defined like below. The production and delivery of quality knowledge and intelligence made of collected information through process of data by clever information analyst.

The Staff-CIO must have
1. Proper education of Humanities and Sciences.
2. Continuous professional training on Art of War.
3. Certification programs of Human Factors.
4. Supplemental education on Informatics and IT.

After the candidate who earned master degree of MBA course, the JIT (Job Instructor Training) is recommended to the new employee. In this training method the instructor first tells trainee what he is to do, and shows him (actually performing task himself), then has the employee repeat the instructions in his own words and watches him and corrects him as he follows instructions, finally repeats the steps until the employee is letter-perfect. One another way is role-playing. Face-to-face contact is vital for training of the employee. This kind of training is generally carried out before the group. This training including group discussions, lectures, and special courses are done at nearby schools or universities.

Then, we may see in future again a modern Staff-CIO, Omura Masujiro.

5 Refference

[1] Hansen,M.T.;Nohria, N; Tierney, T. (1999), What your strategy for managing knowledge?, Harvard Business Review, Vol.77, No.2, PP.106-116.
[2] Simon, H.A. (1976), Administrative Behavior(Third ed), Free Press.
[3] Nonaka, I.; Takeuchi, H. (1995), The Knowledge Creating Company, Oxford Universities Press.
[4] P. H. Sullivan, Value Driven Intellectual Capital, John Wiley & Sons, Inc., (2000)
[5] ToshihikoMiyata,Kibinomakibi, Jinbutusousyo,Yoshikawakoubunkan(2008)
[6] KihachiroKimura,OmuraMasujiro-Den, Choeisya,(2010)
[7] MunehayaTaga, Eisai,Jinbutusousyo, Yoshikawakoubunkan(1998)

Keynote Speech:
Dr. Kazuhiro Kushima
(NTT Laboratories, Japan)

IWIN2011

NTT

Fixed Mobile Convergence
- Current Services and Research Trends -

Sep. 21th 2011

Kazuhiko Kushima
NTT Service Integration Laboratories

NTT

Thank you for your support to Japan earthquake!!

We're exerting ourselves for restoration, too….

2

Restored circumstances(Japan earthquake)   NTT

•Quick restration:
It was restored mostly at the end of April.

385
buildings

4900
base stations

Damaged a building

Fixed Line | Mobile

Service stop building

46
buildings

307
base stations

Damaged telephone poles

Fixed Line | Mobile

4
buildings

27
base stations

18
base stations

Fixed Line | Mobile

0
buildings Mobile

Dengerous Area
※Around Fukushima nuclear plant etc.

9 | 68

Fixed Line 5 | Mobile 17

5 | Mobile 17

Near 11-March
(After the quake hit)

28-March

26-April

30-April

Copyright© 2011 Nippon Telegraph and Telephone Corporation   3

Securing a means of communication   NTT

➢ Installation of a specially-installed public telephone by portable satellite equipment（About 3600 sets)

➢ Installation of movement base station cars(31 cars)

➢ Rent of free satellite cellular phones(About 900 phones)

➢ Rent of free cellular phones(About 2100 phones)

➢ Establishment of free internet use sections(336 locations)

➢ Rent of free tablets (About 670 tablets)

➢ Transmission of living support guide to the earthquake victims on our portal site

Copyright© 2011 Nippon Telegraph and Telephone Corporation   4

- Contents -

NTT

1 Telecommunication Market in Japan

2 Evolution of Fixed and Mobile Broadband Networks

3 Current Status of FMC in Japan (NTT)

4 Research Trends for Future FMC

---

1 Telecommunication Market in Japan

Telecom Service Subscriptions in Japan

NTT

(millions of subscribers)

Mobile subscribers
119.18 M (Sep. 2010)

3G mobile subscribers
106.17 M (Dec. 2009)

Fixed-line subscribers
have declined since 1998

Telephone/ISDN subscribers
36.22 M (Sep. 2010)

Broadband subscribers
34.02M (Sep. 2010)
[Penetration rate for households: 63.8%]

Mobile Phone   Telephone ISDN   3G Mobile Phone   Broadbaod
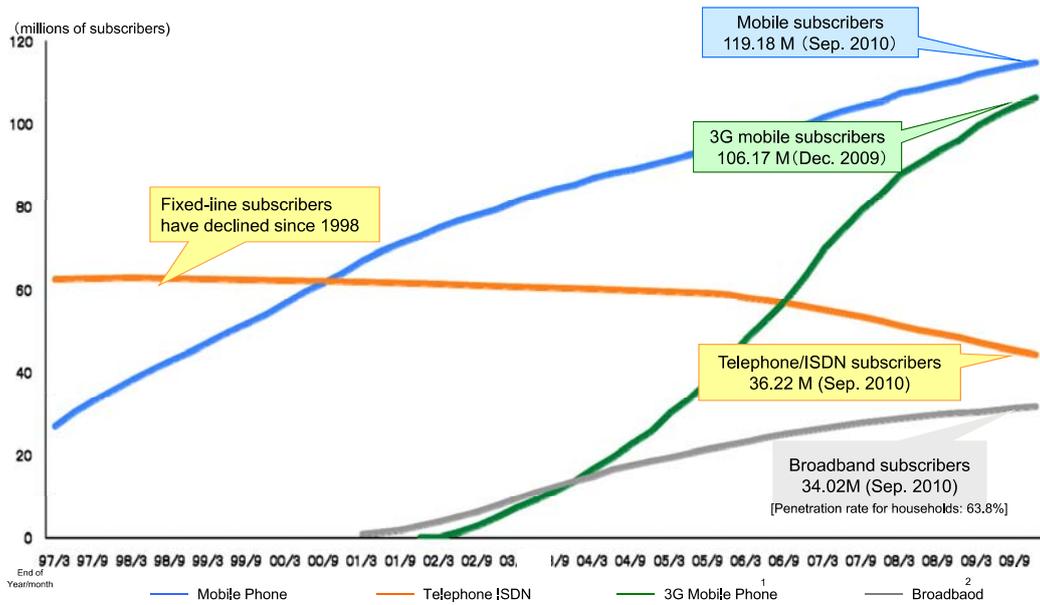
Source: Ministry of Internal Affairs and Communications (MIC) and the Telecom Carriers Association
Note 1   3G mobile phones include W-CDMA and CDMA2000 types.
       2   Broadband includes DSL, FTTH and CATV.

6

1 Telecommunication Market in Japan
Broadband Market in Japan

NTT

➢ Upward trend in NTT East and NTT West's broadband market shares
➢ FTTH market is growing, including transition of subscribers from DSL

(thousands of subscribers)
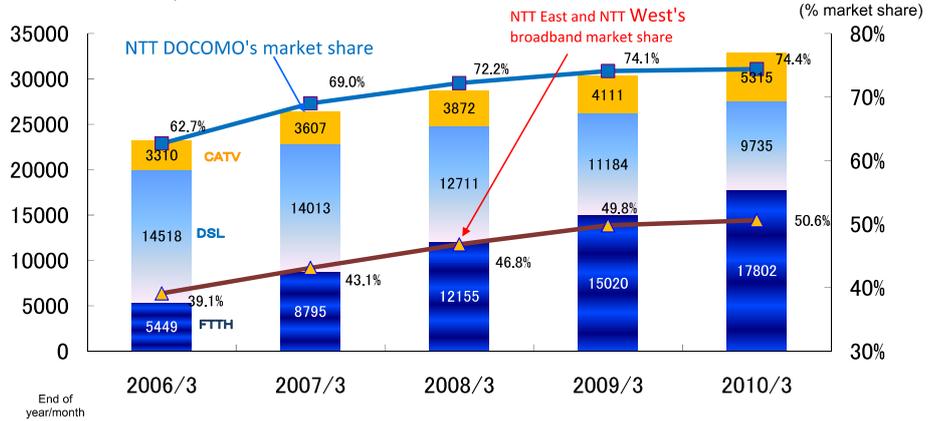
(% market share)

End of year/month

(Source) Ministry of Internal Affairs and Communications

1 Telecommunication Market in Japan
Mobile Market in Japan

NTT

➢ Mobile phone market is gradually expanding
➢ NTT DOCOMO's market share has gradually stopped declining

(thousands of subscribers)

(% market share)

End of year/month

(Source) Ministry of Internal Affairs and Communications

– Contents –

NTT

1 Telecommunication Market in Japan

2 Evolution of Fixed and Mobile Broadband Networks

3 Current Status of FMC in Japan （NTT）

4 Research Trends for Future FMC

---

2 Evolution of Fixed and Mobile Broadband Networks

Change of Fixed Networks (after 1999)

NTT

・ After reorganization of NTT in 1999, NTT East and West changed networks from PSTN to IP by constructing regional IP networks and developed unique FLET'S service to cope with increase of Internet access.
・ Development of regional IP networks is challenge for "communication between prefectures", "flat pricing", and "becoming ISP", exploiting "extended business＊" permitted by revised NTT law.
・ During this period, main technology used in access line has changed from dial-up (PSTN) → ADSL → optical (MC, STM-PON → B-PON → GE-PON). After assignment of IP-phone numbers in 2002, 050 and OA-J IP-phone services have been developed.

＊Enlarged rule defined by second term of revised "NTT law". Assuming that "smooth provisioning of regional communication" and "faire competition" are satisfied, NTT is permitted to provide services with utilizing management resources by being agreed by Minister for Internal Affairs and Communications, in addition to phone service.

2000　　　　　　　　　　　　　　　　　　2005　　　　　　　　　　　2010

▲1999 NTT reorganization
▲2001 Extended business e-Japan
▲2002 Assigning IP-phone number
　▲Resonant framework
▲2005 Midterm management plan "30 million optical access"
▲2008 New midterm management plan "Service creation group"

"First, flat-price phone" (limited time zone and dst. number)

▲2001 AP navi

Fixed price for limited dst. Number by loop-back at ISM

▲1995 Tele-hodai

▲2000 FLET'S ISDN

Germ of IPSP service

ADSL

Dial-up

▲2001 FLET'S square

▲2002 B FLET'S new family （NTT East） B FLET'S family 100 type （NTT West） B-PON

FTTH

Start NGN

▲2000 FLET'S ADSL

▲2004 B FLET'S hyper family （NTT East） GE-PON

▲1999 Tokyo metallic starting ADSL service

▲2001 B FLET'S

▲2001:Y! BB starting ADSL

▲2005 FLET'S Hikari premium （NTT West）

▲2001:Y! BB starting optical service

Communication between prefectures by backnbone operators (Non-extended business)

▲2001 L mode (Connecting AP of each prefecture and GW)

▲2003 Broadband in regional IP network-AP

▲2006 Connecting IPv6 of NTT East and West

▲2008 NGN extended business

Limited services IP connection between NTT East and West

Communications between prefectural IP networks for improved services → Increases ISP count and permits FLET'S Office over prefectures

2 Evolution of Fixed and Mobile Broadband Networks

Hikari/NGN Service Rollout - Service Rollout Centering on NGN

NTT

> **For business users**: Provide highly reliable services, such as FLET'S VPN Wide, that can be used throughout Japan by connecting networks of NTT East and NTT West. Deploy secure bandwidth data transmission services and other means to provide variety of NGN services.
> **For home users:** Increase video services and terminal device lineups that expand use of fiber-optics. Going forward, provide convenient services that combine PCs, TVs, and mobile phones.

		~ FY2008	FY2009 ~

For Business

Provide highly reliable corporate data networks
- Mar: FLET'S VPN Gate
- Mar: Business Ether-Wide
- Mar: Hikari Denwa*1 Office Type
- Aug: FLET'S VPN Wide (Connecting NTT East and NTT West and existing FLET'S service)
- Oct: New key telephone system (compatible with the NGN)
- Present — Secured bandwidth data transmission

ICT solutions that support business creation and efficiency
- Jun: Ubiquitous Desk (NEC)
- Nov: Digital signage (business model trial)
- Oct: Tele presence (Cisco Systems)
- Apr: EMR for clinics
- Jul: SaaS common platform
- Jul: Digital Cinema
- Apr: HD quality video conference system (NEC)

For Home

Provide high-quality triple play service
- Mar: FLET'S Hikari Next
- Mar: Hikari Denwa*1
- Mar: FLET'S Phone
- Mar: Hikari TV
- Oct: Session Plus
- Oct: Session Plus
- Feb: Hikari softphone
- Feb: Hikari digital photo frame
- Jun: Hikari Box (PC that connects the TV set to access the Internet)
- Standby display with communication terminal*2 (Advanced digital photo frame)

Applications and terminal devices that expand use of fiber-optics
- May: Retransmission of DTT over IP (Hikari TV)
- Apr: Earthquake early warning
- Jul: FLET'S TV
- Sep: GyaO NEXT (USEN)
- Dec: "NHK On-demand" (via Hikari TV)
- Aug: Music broadcasting (USEN)
- Apr: PC equipped with Hikari TV tuner
- Sept: Femtocell "My Area"
- Hikari TV and mobile seamless service

Note: Black: network services, Blue: terminal devices and systems. Green: applications and solutions
Italics: services provided by non-NTT Group service providers and manufacturers. (Company names are in parentheses)
*1 Fiber-optic (Hikari) based IP telephony service provided by NTT East and NTT West
*2 Advanced digital photo frame which automatically updates contents
Copyright© 2011 Nippon Telegraph and Telephone Corporation

---

2 Evolution of Fixed and Mobile Broadband Networks

SmartPhone Market (shipped/forecast) in Japan

NTT
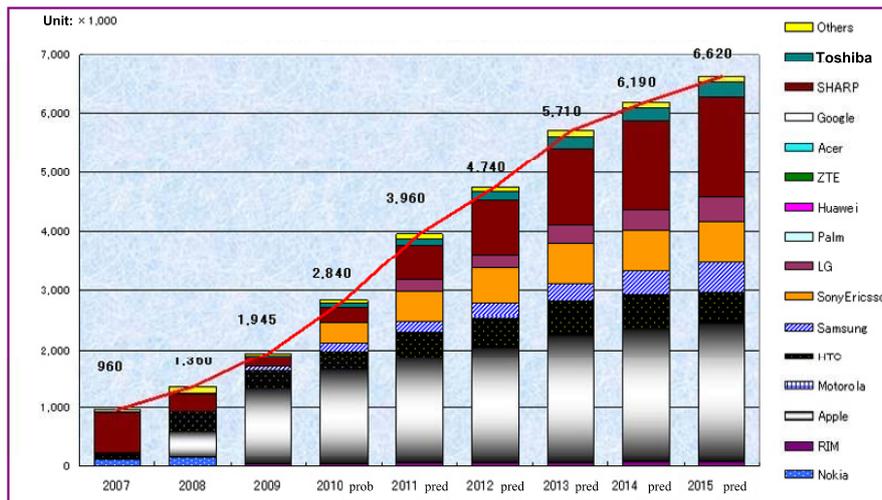
SmartPhone Market in japan: list by terminal vendors (shipped/forecast by YANO Research)

Copyright© 2011 Nippon Telegraph and Telephone Corporation 12

2 Evolution of Fixed and Mobile Broadband Networks

Further Advancement of Handsets

NTT

- Plan to incorporate i-mode mail capability in smartphones, which has been strongly requested by users, by mid-FY2010
- Plan to introduce content billing capability as next step in evolution of "docomo market" by Mar, 31 2011 ⇒ Aim to further expand sphere of smartphone market

| Past | FY2010 | Future |

i-mode handset

Global apps | Operator Pack
Linux
LSI (Application CPU)

N-01B N-02B

Models equipped with Operator Pack
(2009-10 winter/spring model)

Global apps | Operator Pack
Symbian
LSI (Application CPU)

Smartphone

BlackBerry Bold T-01A HT-03A SC-01B

Strengthen sales aiming to acquire 50% share of smartphone market in FY2012

Integrate application platform

Popular i-mode services

XperiaTM (Released Apr. 2010)

"docomo market" (Launched Apr. 2010)

Incorporate i-mode services

i-mode mail support (Xperia)

Develop smartphones compatible with "Osaifu-Keitai" e-wallet service

Introduce content billing capability

Global apps | Operator Pack
Linux/Symbian
LSI (Application CPU)

Smarter product lineup

◆ Xperia is a trademark or registered trademark of Sony Ericsson Mobile Communications AB

13

---

2 Evolution of Fixed and Mobile Broadband Networks

Support of Service Convergence in Cross-industry Fields

NTT

➢ In cross-industry fields between telecommunication industry and industries, such as logistics, transport, and energy, service convergence is advancing
➢ Mobile communication modules are being incorporated into industrial machinery and equipment for data transmission and remote operation of machinery
➢ Going forward, NTT Group will support service convergence for consumer use such as in automobiles and home appliances

Example of service convergence using mobile communication modules

Operation management of construction equipment

Settlement and open space information management at coin parking lots

Freight vehicle delivery management

Taxi/bus operation management

Service convergence in cross-industry fields (logistics, transport, energy, etc.)

Remote monitoring of greenhouses, etc.

Remote reading of water/gas/electricity meters

Settlement and inventory management of vending machines

Settlement using mobile terminals

Expansion to consumer use (ubiquitous)

Expansion in industry

Expansion to consumer use (home)

Content distribution etc. to car navigation systems

Cars

Management of vehicle status, tracking theft, etc.

Remote monitoring of medical equipment

Digital signage

Remote monitoring of elevators

Sending photos

Home Electronics

Sending photos/video from digital or video cameras

Mobile communication modules (NTT DOCOMO): 1.60 million subscriptions (Mar. 2010)

14

2 Evolution of Fixed and Mobile Broadband Networks

Evolution of Networks in Mobile Communications Business

➢ Steadily advanced networks that sustains DOCOMO's video and various other services to provide "higher speeds" and "larger capacity"
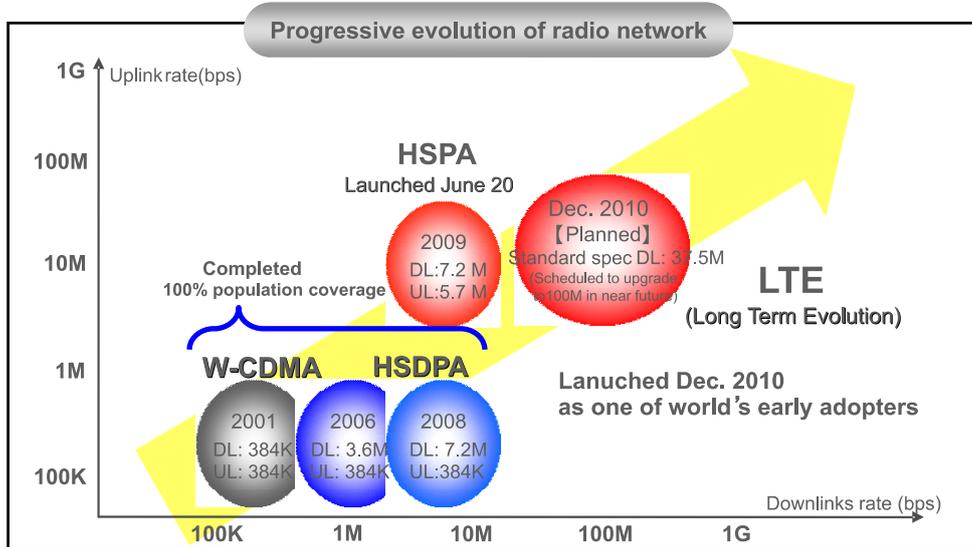
15

- Contents -

3 Current Status of FMC in Japan (NTT)

Fixed-Mobile Convergence (FMC)

NTT

For Business Users	For Home and Consumers
■ **"Mobile Connect"** Access VPN through mobiles or PCs remotely and use e-mail and business applications as in office.	■ **"My Area"** Access movie files or monitor if person is home, by connecting Femtocell base station to FTTH line.
■ **"Biz Communicator"** Remotely access office PCs using fixed/mobile broadband to send e-mail and work as if users were in office. Teleconferencing is also available.	■ **"Hikari Portable"** Wireless LAN router which allows connection to FTTH service, wireless LAN, and mobile phones.
■ **"Biz Mail"** High-quality, high-security, and high-capacity e-mail service based on cloud computing, which may be used on office PCs and mobile phones including smart phones.	■ **"Fan+(plus)"** Content distribution service based on cloud computing, which allows hybrid content composed as combination of texts, pictures, movies, and music to be used on multiple devices including PCs and mobile phones.
■ **"Business mopera IP Centrex"** Enable mobile phone service outside office and IP phone service within office using dual FOMA/Wireless LAN mobile terminal	■ **"DOCOMO Web Mail powered by goo"** E-mail service which may be accessed from PC and mobile phone web browsers.
■ **"IP Centrex One Number"** Single phone-number (090) for both inside and outside office for incoming and outgoing calls	■ **"Home U"** Enable mobile services outside home, while enabling high-speed i-mode and IP phone services by connecting to broadband network inside home, using FOMA/Wireless LAN dual terminal. One-number services also available.
■ **Video telephony communication between FOMA and fixed-line (IP phone)**	■ **Video telephony communication between FOMA and fixed-line (IP phone)**
■ **One Billing for fixed and mobile (consolidated billing)**	■ **"goo One Billing"** single screen display for fixed and mobile billing

(vertical text: Fixed-Mobile Convergence Services)

17

---

3 Current Status of FMC in Japan (NTT)

Keitai Home System: terminal-linkage-type FMC

NTT

■ In Sep. 2009, NTT docomo announced development of "Keitai Home System" and provides it to home developers and homemakers as solution for controlling home appliances.

■ "Keitai Home System" enables one to check visitors with video by TV intercom linked with mobile phone using controller for home appliances, e.g., air conditioner and lights, and provide information to mobile phone.

Inter-phone linkage
House appliance (HA) control
Electric key linkage
HA adapter
Sensor
Wireless LAN access point
Electric Key
Inter-phone (master)
Network camera
Home controller
Communication by wireless (ZigBee)
Public network
Broadband router
Internet
Collaborate with inter-phone controlling HA from outside
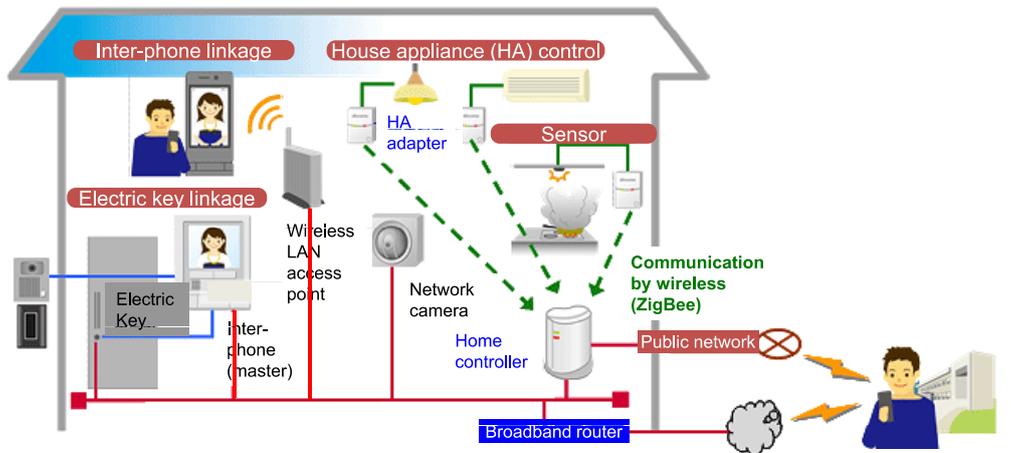
From web page of NTT docomo,
(http://www.nttdocomo.co.jp/info/news_release/page/090924_01.html#p01)

3 Current Status of FMC in Japan (NTT)

Personal Wireless Router (PWR)

NTT

■**PWR, mobile WiFi router*1, enables one to use portable game devices and PCs, which have been used in WiFi area, in areas where public wireless LAN or 3G networks can be used.**

FOMA network of NTT docomo

📶 **3G**

Public wireless LAN

Optical network

Mobile WiFi router

WAN interface:
3G, wireless LAN, Ethernet
Max. number of terminals: 6
Max. communication time: 6 hours

Potable game   PC   Digital camera   Digital music player   Tablet terminal
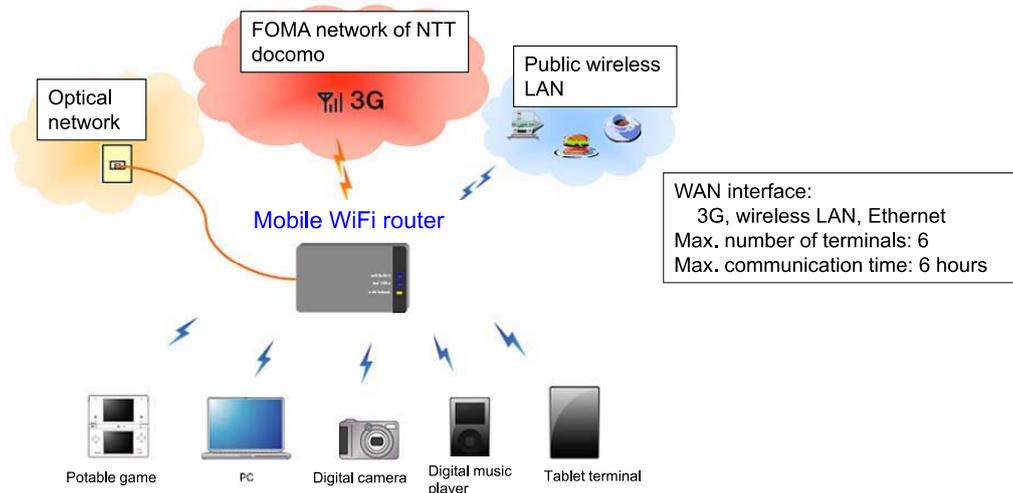
*1 Mobile WiFi router: cooperative development by NTT broadband platform and buffalo Inc.

From web site of NTT docomo, http://www.nttdocomo.co.jp/info/news_release/2010/05/18_01.html#p04

– Contents –

NTT

1 Telecommunication Market in Japan

2 Evolution of Fixed and Mobile Broadband Networks

3 Current Status of FMC in Japan (NTT)

4 Research Trends for Future FMC

4 Research Trends for Future FMC (NTT)

Integration Structure and Current Target of FMC

NTT

■ITU-T recommendation（appendix of Y.2802/FMC）defines four integration scenarios of FMC, shown in blue columns.
■Demos for investigating service concept and experimental development for （i）remote collaboration service and （ii）interactive digital signage as AP layer integration.

※From appendix of ITU-T Y.2802（requirement condition of FMC）

Non FMC（separate NW）	Control layer Integration	AP layer integration (1)	AP layer integration (2)	Control layer integration
Charging function / Charging function Application / Application ASF&SSF / ASF&SSF Service control / Service control Fixed NW / Mobile NW Fixed access / Mobile access	Charging function Application / Application ASF&SSF / ASF&SSF Service control / Service control Fixed access / Mobile access	Charging function Application ASF&SSF / ASF&SSF Service control / Service control Fixed access / Mobile access	Charging function Application ASF&SSF Service control / Service control Fixed access / Mobile access	Charging function Application ASF&SSF Service control Fixed access / Mobile access
Independent NW func. Service integration by shared terminal	Can merge all charges to single bill	Can use application with shared form	Can use application support function	Integrate core NW and all functions
Difficulty	Easy ←		→ Hard	

Copyright© 2011 Nippon Telegraph and Telephone Corporation

---

4 Research Trends for Future FMC (NTT)

(i) Field Work Assistance by Remote Collaboration

NTT

■ Remote collaboration through videophone between a support center on the fixed side and a customer or on-site technician on the mobile side is assisted by overlaying graphics （AiRstamps） on on-site video image. AiRstamps enable direct instructions that are difficult to convey by words alone making the collaboration more efficient.
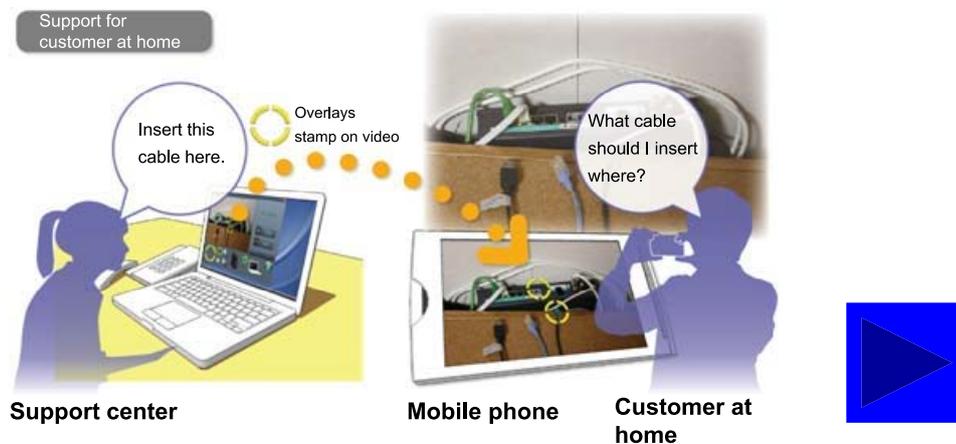
Support for customer at home

Insert this cable here.

Overlays stamp on video

What cable should I insert where?

Support center　　　　**Mobile phone**　　**Customer at home**

Copyright© 2011 Nippon Telegraph and Telephone Corporation

4 Research Trends for Future FMC (NTT)

(i) Field Work Assistance by Remote Collaboration (cont.) NTT ◑

- **Remote Collaboration Service is a combination of FMC (Fixed Mobile Convergency) video communication and AR technology, which supports onsite workers remotely.**
- **Regardless of the mobile phone's movement such as jiggling, pan tilt, or putting it aside, "AiR Stamp" can be always attached on the target object with AR Technology.**
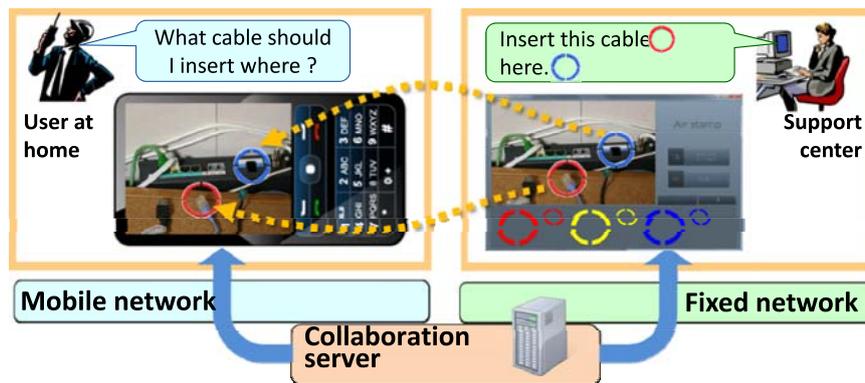
4 Research Trends for Future FMC (NTT)

(ii) Interactive Digital Signage with Mobile Phones NTT ◑

- **Town advertising services that temporarily link and synchronize mobile phones and digital signage.**
- **On seeing an ad displayed on digital signage, the user can get more information by dialing the number for the corresponding voice guidance. This triggers the digital signage to display video for that user while voice guidance synchronized with the video is output from the user's mobile phone.**

Digital Signage Display & Mobile Phone

1	2	3	4
Watching ads	Making a call	Operating With guidance	Interactive voice & visual response

4 Research Trends for Future FMC (NTT)

(ii) Interactive Digital Signage with Mobile Phones (cont.) NTT

■ "Interactive Digital Signage" is a new type of town advertising service, characterized by interactive voice response and cooperating visual contents.
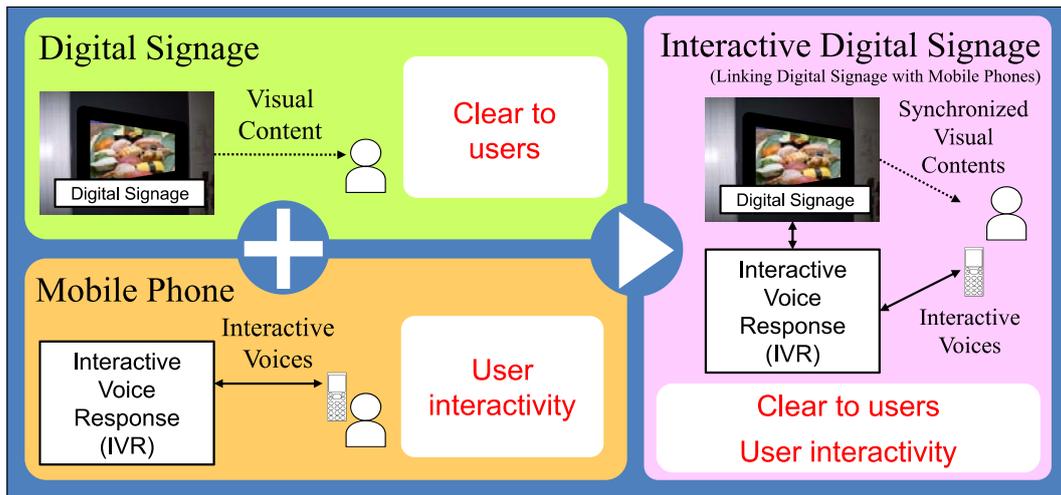
Digital Signage

Visual Content

Clear to users

Digital Signage

➕

Mobile Phone

Interactive Voices

Interactive Voice Response (IVR)

User interactivity

Interactive Digital Signage
(Linking Digital Signage with Mobile Phones)

Synchronized Visual Contents

Digital Signage

Interactive Voice Response (IVR)

Interactive Voices

Clear to users
User interactivity

---

4 Research Trends for Future FMC (NTT)

Multi device: combination of multiple device usage NTT

■ It is becoming popular among personal users to use many networked high-functional devices.

■ As a consequence, one service is provided to a user by using multiple devices dependent on the situation.

■ Interactive Digital Signage falls under the "controlling" category.

Controlling	**Sharing**	**Takeover**
Mobile AP*1 Fixed *1: Application	Mobile AP Fixed	Mobile AP Fixed
A terminal controls the other terminal(s).	Networks share resources (contents, user profiles, etc…) of AP.	A user can use AP on fixed/mobile networks continuously.
Mobile AP AP Fixed	Mobile AP Fixed	Mobile AP Fixed Internet Clouds Services
A notification from one AP induces a user to use another AP.	Multiple user(s) interact with each other through both networks.	Services on fixed/mobile networks and Clouds mix. (Mashup)
Inducing	**Interacting**	**Mashup**

Conclusions

NTT

■On active mobile communication and shrinking fixed communication, some overseas operators are reconstructed with merging mobile and fixed networks, and FMC is becoming popular in Japan as well. NTT Group is also planning to provide ubiquitous services using FMC.

■NTT laboratories propose concept of application service for creating new FMC services, investigating techniques for convergence between mobile and fixed networks, and engaging in standardization activities.

Thank you for your attention