NTT DOCOMO

Technical Journal

Vol.20 No.2 Nov. 2018

NTT docomo

Toward Life-enhancing and Secure Services

General Manager of Service Design Department
Takaaki Sato

The Service Design Department is engaged in a wide range of activities, such as the development and implementation of diverse services and applications, the development and construction, maintenance and operation of i-mode, sp-mode, and other platform systems, the operation of security systems, and the provision of various types of technology support. These services and systems include "+ message" as an evolution of the existing Short Message Service (SMS), "my daiz" as an evolution of the existing i-concier and Shabette Concier personal information services using Artificial Intelligence (AI) technology, and a membership system to support the upgrading and evolution of the business infrastructure and expand commercial distribution of products and services. Such wide-ranging and integrated activities from service development to maintenance and operation is a strong point of the Service Design Department, which aims for speedy development applying the technical knowledge, experience, and strengths cultivated to date, agile development to meet customer needs as quickly as possible, and efficient construction, maintenance, and operation of safe and secure infrastructures making aggressive use of advanced technologies. As recent examples of the work being performed in the Service Design Department, I would like to introduce 1) efforts toward the practical use of "Mieru Denwa®*1" and 2) large-scale introduction of a virtual platform for treasure Casket of i-mode service, high Reliability platform for CUStomer (CiRCUS)*2/Multi Access Platform System (MAPS)*3.

1) Efforts toward the Practical Use of Mieru Denwa

Mieru Denwa is a voice-call support service targeting hearing-impaired persons who have difficulty in using voice calls. It can convert the spoken words of the other party into text in real time and display them on a smartphone screen. The provisioning of convenient functions that can support the handling of calls and enhance the lives of hearing-impaired persons has been an issue of special concern for telephone operators.

This service has been made easy-to-use for hard-of-hearing persons by starting with "voice recognition technology" developed over many years by NTT DOCOMO and adding a "speech support function" that reads input text and conveys that content to the other party. Furthermore, to build a bridge between "text-based communication" and "voice-based communication" and achieve smooth communication, Mieru Denwa adds special features such as displaying animation to indicate speaking by the other party and text display of the other party's speech in real time.

A strong point of this service is that it can be connected to other services such as "Mieru Rusuden®*4" that enables the user to check a voice message recorded by the call-answering function in text form and "Ohanashi Memo™*5" that converts a phone conversation to text. Strong points like these can be accumulated and generalized for application to a variety of fields.

2) Large-scale Introduction of a Virtual Platform in CiRCUS/MAPS

OpenStack*6 has been introduced into NTT DOCOMO's CiRCUS/MAPS large-scale, mission-critical*7 system with the aim of reducing infrastructure costs and improving response speed in service development. OpenStack is Open Source Software (OSS) that speeds up function development and enables the use of many functions at low cost. However, the characteristics of OSS can also make it difficult to maintain a stable level of quality. We therefore performed the following measures when setting out to adopt OpenStack: (1) conducted three Proof of Concept (PoC)*8 trials before function development and thoroughly extracted bugs, (2) adopted architecture that minimizes the impact of an OpenStack problem on services (excludes a Single Point of Failure (SPOF)*9), and (3) rigorously limited the functions to be used while adopting more stable technologies. As a result of these efforts, CiRCUS/MAPS has been operating without any major problems up to the present.

In this way, we automated work using a virtual platform having extensive application program interfaces (APIs). For example, we have reduced manpower by about 50% and work time by more than 90% for night work associated with End of Life (EOL) migration that in the past extended over several days in a physical environment.

In recognition of this achievement of introducing a virtual platform to CiRCUS/MAPS on a large scale, we received a Red Hat Innovation Awards APAC 2017.

Going forward, the Service Design Department is committed to providing services that enhance our customers' lives and to providing safe and secure network connections. We will expand "proposal-type" service development and technology support making use of accumulated strengths and work to provide a forward-looking service platform.

*1 Mieru Denwa®: A registered trademark of NTT DOCOMO.
*2 CiRCUS: A device that serves as an interface between the NTT DOCOMO core network and the Internet, provides i-mode mail, i-mode menu, ordinary Internet access, and other functions.
*3 MAPS: A platform providing Internet and business-system connections from FOMA, Xi, and other access circuits.
*4 Mieru Rusuden®: A registered trademark of NTT DOCOMO.
*5 Ohanashi Memo™: A trademark or registered trademark of NTT DOCOMO.
*6 OpenStack: Open source software for constructing an IaaS-type cloud-computing environment. It can be used to construct a virtual cloud environment for each user cloud service using physical resources such as servers, storage equipment, and networks.
*7 Mission critical: Refers to elements essential to the execution of businesses or services for which termination or interruption is not allowed.
*8 PoC: A relatively simple demonstration of the significance or feasibility of a new concept or idea.
*9 SPOF: A single point or element or that can cause an entire system to stop operating when failing in a certain part of the system.

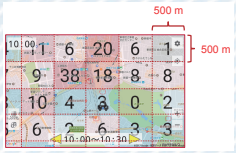
NTT DOCOMO
Technical Journal
Vol.20 No.2 Nov. 2018

[Contents]

DOCOMO Today

Technology Reports (Special Articles)

Special Articles on AI Technologies Contributing to Industry and Society



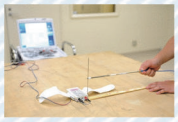
(P.14)

Technology Reports

(P.32)

(P.42)

Collaboration Projects



(P.60)

Investigative Research into Electromagnetic Interference with Medical Equipment from Radio Waves Emitted by Mobile Telephones and Smartphones 60

Healthcare ICT | Electromagnetic Interference | Safety and Security

Topics

(P.68)

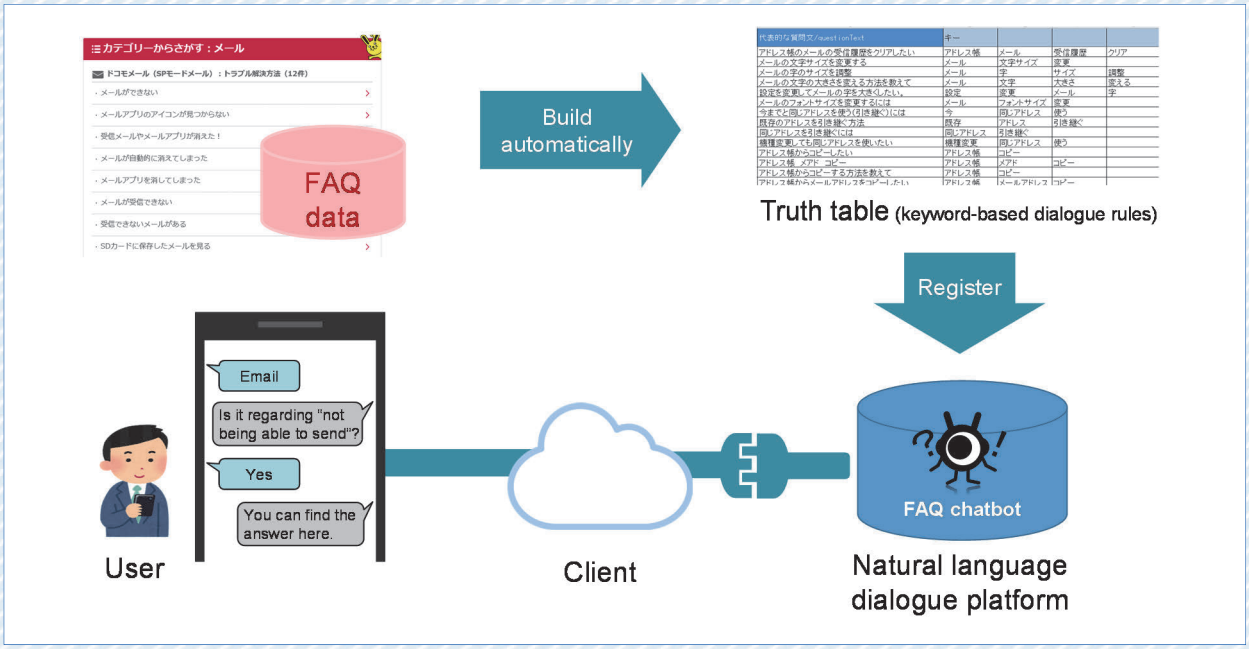
Low-power LTE Terminal for Gas Smart Meters 68

eDRX | LPWA | Smart Meters

News

(P.72)

Medal with Purple Ribbon in 2018 Spring Medals of Honour 72



Technology Reports (Special Articles)　Automatic FAQ Chatbot Building Technology (P.4)

FAQ chatbot overview

Technology Reports | Natural-language Conversation | Chatbots | FAQ
Special Articles on AI Technologies Contributing to Industry and Society

Automatic FAQ Chatbot Building Technology

Service Innovation Department Hiroshi Fujimoto Takanori Hashimoto Yuriko Ozaki
Innovation Management Department Takuya Kobayashi Koji Yamazaki

FAQ chatbots that use AI to provide responses instead of a human are becoming more widespread. In the past, when building such FAQ chatbots, scenarios had to be built-in or incorporated using machine learning, so it was costly to generate each scenario and accuracy of the responses was lower than desired. NTT DOCOMO has developed a technology to generate chatbot response scenarios automatically from a set of FAQ question and answer texts. This enables the FAQ provider to build a chatbot easily, without the need to develop each scenario for the chatbot.

1. Introduction

It has become common practice for enterprise and administration call centers and other support departments to replace parts of their work with chatbots*1 to reduce costs. Frequently Asked Questions (FAQs) repeatedly answer similar questions for different people, so they are particularly well suited to having a chatbot provide responses mechanically.

In the past, chatbots specializing in handling FAQs (hereinafter referred to as "FAQ chatbots") have generally been built based on either a scenario model, or a category model.

For a scenario-based chatbot, scenarios that branch the dialogue with the user are generated based on each of the user's queries, and responses defined for each branch are returned. For an FAQ, the chatbot continues the dialogue, even if the user's query is ambiguous, digging into the query and narrowing the possible responses until it can give an accurate response. However, as the number of

*1 Chatbot: A program that automatically conducts dialog with people with speech or text chat.

branches before reaching a response increases, the cost of generating the scenarios also increases, so this approach is not suitable for large FAQs containing hundreds or thousands of possible responses. Examples of services providing scenario-based chatbots include Repl-AI®*2 [1] and IBM Watson™*3 Conversation [2].

Category-based chatbots select the response that is best suited to the user's query from among prepared responses and present it to the user. Generally, the algorithm used to select an appropriate response is supervised machine learning*4. A single response is given to the user's query, so there is no need to prepare branching scenarios as with scenario-based chatbots, and it is relatively easy to apply this method to large-scale FAQs. However, it is not possible to clarify a user's query with on-going dialogue if it is ambiguous, so the accuracy of responses is generally lower than with scenario-based chatbots. One example of a service providing category-based chatbots is Microsoft Azure™*5 Bot Service [3].

NTT DOCOMO has developed an entirely new, automatic FAQ chatbot building technology to resolve issues with the scenario-based and category-based techniques described above. The proposed technology supports on-going dialogue through scenario branching and is able to generate scenarios for each branch automatically using a newly developed proprietary algorithm. Thus it is able to maintain the accuracy of a scenario-based chatbot, while being applicable to large-scale FAQs, like a category-based chatbot. Note that the QnA Maker API [4] from Microsoft is an existing technology for building FAQ chatbots automatically, but it is only able to build category-based chatbots, and to the best of our knowledge, no other technology able to build scenario-based chatbots automatically currently exists.

NTT DOCOMO has also developed an automatic FAQ chatbot building system using the proposed technology. The system is a Web application that operates by calling an API implementing the proposed technology, and provides functionality including a GUI for building FAQ chatbots, visualization of user dialogue history, and links with existing chat platforms such as LINE*6 and Facebook Messenger®*7. The system has been providing services for corporate enterprises since March 2018.

This article describes the automatic FAQ chatbot building technology, and an evaluation of its performance. It also describes the automatic FAQ chatbot building system that we have developed.

2. FAQ Chatbot

2.1 Overview

An overview of the FAQ chatbot is shown in **Figure 1**. With this technology, intermediate data called a truth table is first generated from the FAQ data. A scenario-based FAQ chatbot is then built by registering the truth table in a natural-language dialogue platform [5] developed by NTT DOCOMO.

1) FAQ Chatbot Dialogue

An example of dialogue with the FAQ chatbot is shown in **Figure 2**.

The FAQ chatbot operates assuming that pairs of queries and responses, referred to as FAQ data, has been prepared ahead of time. Thus, by finding a query that matches the user's intent, a corresponding response can also be identified.

Next, we discuss methods for correctly identi-

*2 Repl-AI®: A registered trademark of Intermedia Planning, Inc.
*3 IBM Watson™: A registered trademark of International Business Machines Corp. in many countries around the world.
*4 Supervised machine learning: A general term for algorithms that learn to recognize patterns in unknown data through repeated pattern learning based on training data.
*5 Microsoft Azure™: A registered trademark or trademark of Microsoft Corp. in the United States and other countries.
*6 LINE: A trademark or registered trademark of LINE Corp.
*7 Facebook Messenger®: Facebook and Messenger are trademarks or registered trademarks of Facebook, Inc.

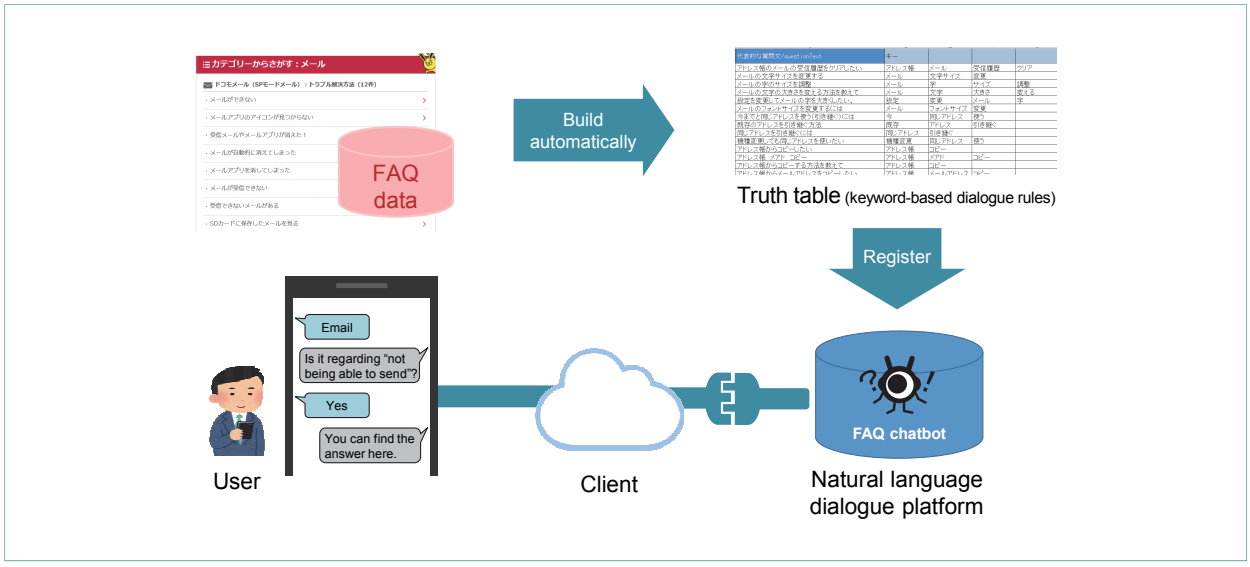


Figure 1 FAQ chatbot overview

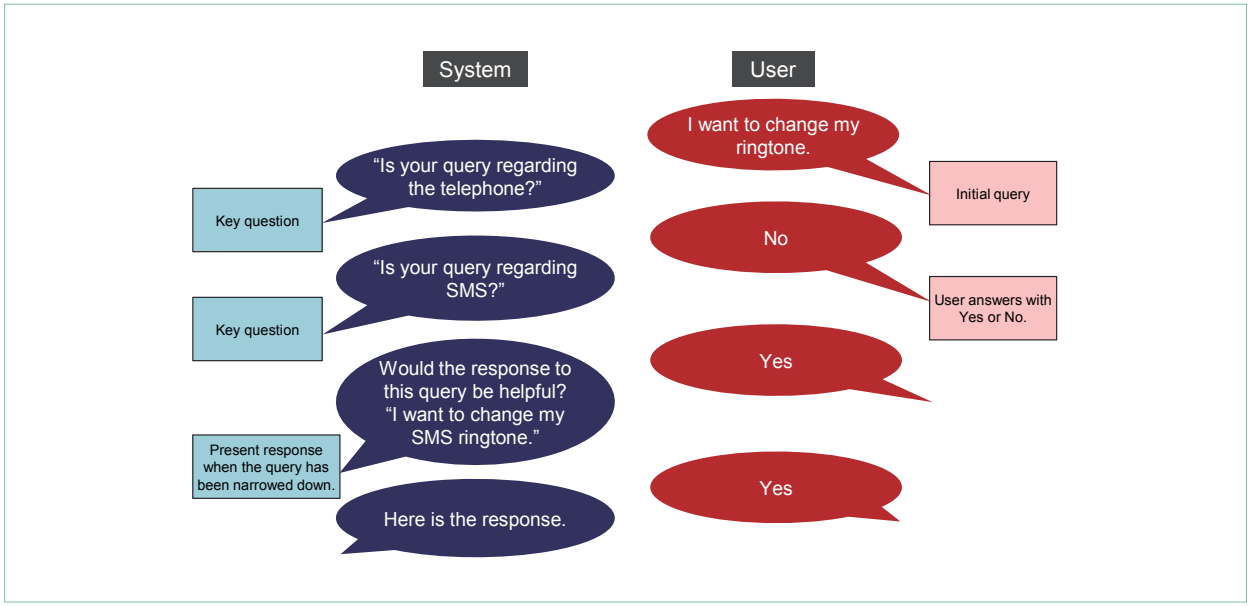


Figure 2 Example interaction

fying queries in the FAQ chatbot. Often, a user's first input ("Initial query" in Fig. 2) is ambiguous, so it can be difficult to identify a query text from just the initial query. In such cases the user is asked additional questions ("Key question" in Fig. 2) to update the query information, and ultimately find an applicable query text.

2) Identifying Queries Based on the Truth Table

The FAQ chatbot searches for queries based on the truth table, which is used to generate dialogue

scenarios automatically, according to the user's query. It is composed of a list of query texts together with key phrases from the query texts, hereinafter referred to as "keys," as shown in **Figure 3**. Keys are phrases that a user is predicted to input when searching by keyword search for an applicable FAQ entry. The proposed system generates them automatically from the query texts. Since there are various expressions for a given key that users could potentially use, in addition to the truth table, the proposed technology also generates a variation dictionary, which stores expression variations for each key.

The FAQ chatbot searches the truth table based on the user's initial query to identify the most applicable query text. If it is difficult to identify a query text from the user's initial query, the system asks key questions to narrow down the possible query texts. Details of FAQ chatbot operation are described below.

2.2 Generating Truth Tables and Variation Dictionaries

To build an FAQ chatbot, the proposed technology generates a truth table and a variation dictionary, as described below.

1) Generating the Truth Table

The truth table is a collection of query texts broken down into keys, which are essential elements of each question (Fig. 3). The keys are not simply a collection of morphemes[*8], they are selected according to a particular algorithm in the proposed technology. Generating the truth table involves two processes: extracting keys using parts-of-speech rules, and discarding unnecessary keys based on a stop key list.

For example, the query, "I don't know my Wi-Fi®[*9] password" is separated into the keys, "Wi-Fi," "password," and "I don't know" (Q1 in Fig. 3). From the results of morpheme analysis, phrases that do not have meaning on their own, such as prepositions and helper verbs, are discarded, essentially selecting verbs and nouns as keys. However, if just single nouns and verbs are selected, special terms with multiple nouns, such as "docomo Wi-Fi," would be separated into the keys, "docomo" and "Wi-Fi." To resolve this sort of issue, we use combining rules to combine sequential nouns and numbers into a single key (Q2 in Fig. 3).

There are also words that are not suitable as keys even though they are nouns or verbs. For

	Query	Key		
Q1	I don't know my Wi-Fi password.	Wi-Fi	Password	~~Don't know how~~
Q2	I want to connect to docomo Wi-Fi.	docomo Wi-Fi	Connect	Stop key
Q3	How do I send email?	Email	How to send	~~Want to know~~
Q4	Sending email	Email	Sending	Stop key
Q5	I can't delete photos.	Photos	Cannot delete	Delete_incomplete
Q6	I accidentally deleted a photo.	Photos	Deleted	Delete_complete

Combine

Figure 3 Truth table example

*8 Morpheme: When dividing-up sentences or phrases, the smallest unit that carries meaning. Involves parts of speech such as nouns and verbs.

*9 Wi-Fi®: A registered trademark of the Wi-Fi Alliance.

example, the query, "How do I send an email?" is separated into "email," "send," and "How do I," but "How do I" is used so often in a FAQ that it is not helpful as a key to distinguish among queries. To remove such unnecessary keys that the parts-of-speech rules have created, key removal rules called "stop keys" are used (Q1, Q3 in Fig. 3). The stop keys are a list created by using the FAQ data to select keys that are used often in the genre, such as "PC," "finance," or "tourism," together with synonyms for those keys.

In this way, only keys suitable for the FAQ chatbot are selected to generate the truth table.

2) Variation Dictionaries

Variation dictionaries are broadly categorized into those that absorb variation in terminology specific to each FAQ and those that absorb variation in more general phrases.

The former lists variations in representation of specialized terminology used in the prepared query texts, and are registered at the discretion of the FAQ chatbot designer. The latter are generated using the large volume of language resources in the natural language interaction platform, and apply to keys with variation in phrasing or usage, such as verbs and action nouns.

Variations in general phrases includes the differences between verbs and action nouns, as well as differences in conjugated forms of verbs. Depending on the query, verbs and action nouns, such as "transmit" and "transmission," mean the same thing, so one key is used. If there was no function to absorb this sort of variation, a query with "transmit the email" would not match a query with "transmission of the email." The variation dictionary absorbs such general variation based on a synonym dictionary with approximately 1,400 verb/action nouns.

Variation due to verb conjugations are absorbed based on patterns of variation in usage for each verb, collected from analyzing approximately 400,000 FAQ data items. Specifically, all verb conjugation patterns are normalized using the expression-normalization function of the natural-language dialogue platform [5], reducing them to the three categories of positive, negative, and completed. These three categories are the most used in FAQs. Negative expressions like "I cannot delete a photo," and completed expressions like "I accidentally deleted a photo," in particular, affect the response that needs to be presented, so they are managed as separate categories (Q5, Q6 in Fig. 3).

3. FAQ Chatbot Behavior

The basic operation of the FAQ chatbot system is shown in **Figure 4**.

The FAQ chatbot repeats the following to refine a query.

- Extract keys from the text of the user's query and update the search key list ((1) to (3) in Fig. 4).
- Search the FAQ and decide on the response content using the search key list ((4) to (5) in Fig. 4.

Each of these behaviors is described in detail below.

1) Key Extraction and Search-key List Update

The FAQ chatbot uses the phrase matching function of the natural-language dialogue platform to extract the keys from the user's initial query

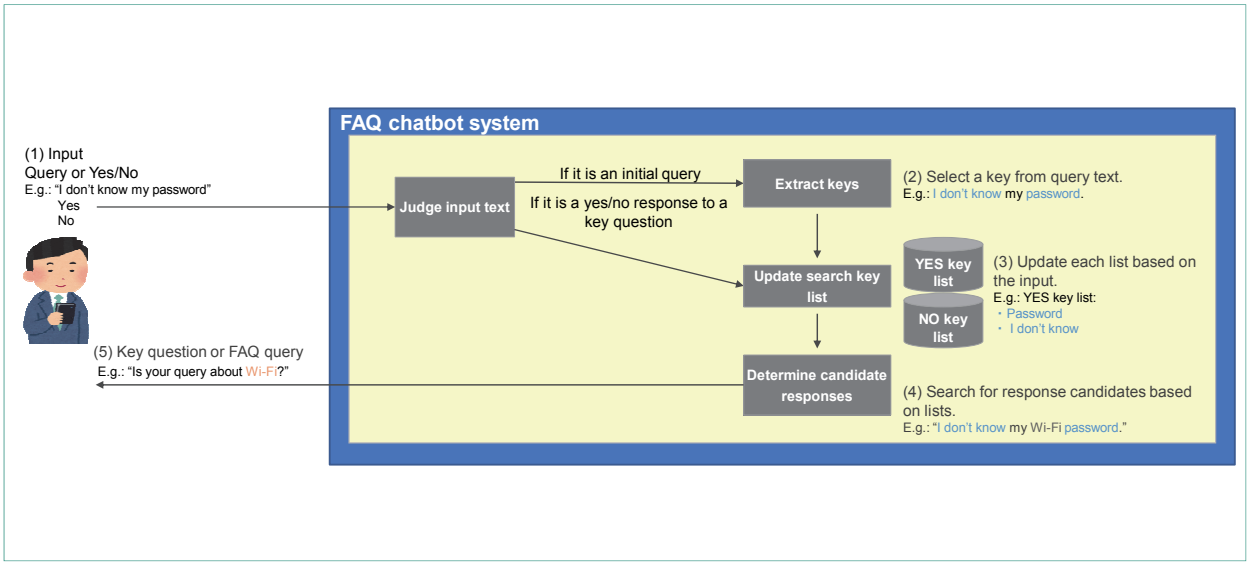


Figure 4 Basic FAQ chatbot system operation

text and matches them with the set of keys for each query in the truth table. In doing so, the user is not expected to be aware of which keys are actually in the truth-table queries, but the variation dictionary is used to absorb some of the variation in key expressions.

The search key list is then updated based on the keys obtained from the user's utterance. The FAQ chatbot keeps two search key lists: a "YES key list" of keys associated with a positive intention, and a "NO key list" of keys associated with a negative intention. A query is identified by searching the truth table with these lists. The YES key list includes keys from the user's initial query and updates from key questions that the user has answered in the affirmative. The NO key list includes updates from key questions that the user answered in the negative.

2) FAQ Search Using the Search Key Lists and Deciding the Response

The FAQ chatbot uses the two search key lists described above to search for queries stored in the truth table. Specifically, it searches for queries that have keys registered in the YES key list and do not have keys registered in the NO key list. Queries may have multiple matches, and such cases are categorized as follows.

- Complete match: A query for which all keys in the truth table entry match with the YES key list, and none match with the NO key list.
- Incomplete match: A query for which only some of the keys in the truth table entry match the YES key list, but not enough for a complete match.

If all matches for the query are incomplete, the FAQ chatbot replies with a question regarding a search key that was not matched (This is called a key question. For example, the chat bot replies "Are you asking about Wi-Fi?" to see if the key is relevant to the user's intention). If the user answers

affirmatively to a key question, the key is added to the YES key list, otherwise it is added to the NO key list. One or the other list is updated and the FAQ chatbot repeats the search for a query in the truth table.

If a complete match is found, the system asks the user to confirm the query. This is called an FAQ query. For example, something like "Does this query express what you mean?" (followed by the selected query), confirms that it matches the user's intentions. If the user answers in the affirmative, the dialogue is finished, and the answer to the confirmed query is returned. If the user answers negatively, the next FAQ or key question is presented.

4. Performance Evaluation

4.1 Evaluation Conditions

We evaluated the performance of the FAQ chatbot with real users, and discuss the results of evaluating the proposed technology below. The FAQ chatbot was built using FAQ data generated from the manual of a particular internal system, including 306 answer texts.

We measured the accuracy of responses and the average number of interactions to reach a response. Response accuracy was the proportion of initial queries for which the FAQ chatbot was ultimately able to present a correct response, and the average number of interactions was the average number of interactions needed to reach an answer.

After performing the first evaluation, we tuned the FAQ data, made additions to the variation dictionary and stop key list, and performed a second evaluation. The first time, we used the FAQ data as-is, applied the proposed technology and conducted the evaluation, but for the second time, we applied some tuning and then re-evaluated to see how much the accuracy could be improved.

Specific details of the tuning are described next. The FAQ data used for the first evaluation was not adequately complete building the FAQ chatbot, and contained issues such as duplicate queries or answers, and errors in the responses themselves. This resulted in cases where the FAQ chatbot was unable to extract suitable keys, or where it was unable to present a correct answer, even when a query text was identified using key questions. Accordingly, we ensured that there were no duplicate queries for any of the answers, and tuned the queries to be more concise, so that keys could be extracted correctly. We also did not use a variation dictionary or stop key list for the first evaluation, so variations in the keys from user queries were not absorbed and unnecessary keys were selected, resulting in degraded performance. We added key variations to the variation dictionary for these errors, and added appropriate stop keys to deal with these issues.

In practical operation, the cost of this sort of tuning will also need to be considered. To do so, we checked the actual amount of work involved in tuning for the second evaluation. Note that for the first and second evaluations, we used different people conducting the evaluation, to ensure that the same questions were not asked for both evaluations.

4.2 Evaluation Results

The accuracy of answers, and the average number of interactions are shown in **Table 1**.

Table 1 Evaluation results

Evaluation	Response accuracy	Average no. of interactions
First	77%	4.3
Second	91%	1.8

We first discuss the results of the first evaluation. The table shows that the FAQ chatbot was able to respond to user queries with a high accuracy of 77%. In this case, the FAQ chatbot was built using FAQ data with absolutely no tuning. This demonstrates that the proposed technology is able to build a chatbot that can answer queries with nearly 80% accuracy immediately. However, the number of interactions before reaching an answer was 4.3, which is high. This was due particularly to having too few stop keys, which resulted in extraction of unnecessary keys, and subsequently, the generation of wasteful key questions.

Next, we discuss the results of the second evaluation. After tuning, the evaluation resulted in very high response accuracy of 91%. The most effective effort was expansion of the variation dictionary. This corrected errors that prevented the correct response from being presented because the keys in the query text were not in the initial query, so it was not found. Tuning the FAQ data also helped in extracting appropriate keys from the query texts and associating query texts with the correct answers, so accuracy of answers increased. By expanding the stop key list, the average number of interactions was also improved significantly, to 1.8.

The above shows that even higher accuracy can be achieved by tuning the input data.

Lastly, we discuss the amount of work required to perform the tuning. Tuning involved inspection work, analyzing the results from the first evaluation and extracting the causes of degradation in the accuracy; generating and updating FAQ data according to the inspection work; and expanding the variation dictionary and stop key lists according to the inspection work. These tasks required 4.4, 3.3, and 0.7 person-days respectively, totaling a relatively short time of approximately 8 person-days. Note that this resulted in addition of 295 entries to the variation dictionary and 53 entries to the stop key list.

5. Automatic FAQ Chatbot Building System

5.1 Overview

The FAQ chatbot provides an API for data management. The automatic FAQ chatbot building system is a Web application with an FAQ chatbot building function that uses this API. With conventional scenario-based FAQ chatbots, the developer of the chatbot service would need to design chatbot dialogs from the FAQ data to build the chatbot. With this system the developer only needs to upload the FAQ data using a Web browser. The system automatically designs the dialog content and builds the chatbot, so the chatbot can be built in much less time.

5.2 PDCA Cycle to Improve Chatbot Response Accuracy

When considering providing a real service, ongoing improvement in the accuracy of FAQ chatbot responses is a major issue. With our system, building and improvements to the FAQ chatbot can be done at the same time. Improving accuracy of responses is done by repeating the process shown in **Figure 5**. First, the developer downloads the dialog history of the FAQ chatbot and users on the system. Next, the queries that the FAQ chatbot was unable to answer are found, and answers to them are added to the FAQ data. This updated data is then be uploaded through a Web browser, thereby completing update of the FAQ chatbot. Repetition of this update process increases the accuracy of FAQ chatbot's responses. As with building the chatbot, there is no need to design dialogs when performing this type of update. The work is done mechanically, so these improvements can be done quickly.

5.3 Linking with Existing Chatbot Platforms

The ability to build and update FAQ chatbots mechanically and in a short time is an advantage, but responding with text input to repeated questions places a burden on users. To reduce this burden, the system is able to display buttons, as is done by other existing chat platforms. Specifically,

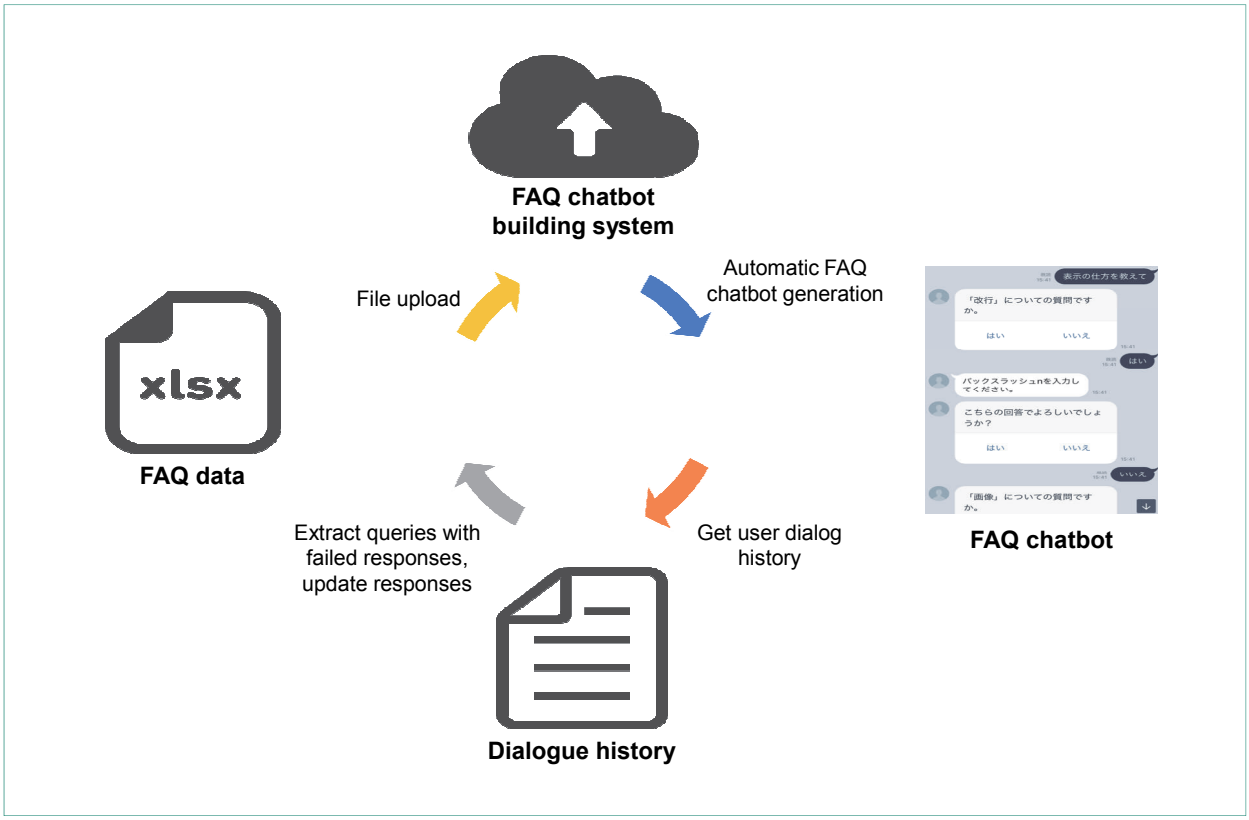


Figure 5 PDCA cycle to improve quality

when the FAQ chatbot asks the user a question, the user can respond by simply pressing a Yes or No button as they make their way to the answer.

6. Conclusion

This article has described a technology to build FAQ chatbots automatically. The proposed technology automatically builds a chatbot from FAQ data, supports processing of large volumes of FAQ data, and is able to respond to user queries very accurately through the use of automatically-generated dialogue scenarios. We evaluated the system using real FAQ data, and the FAQ chatbot achieved accuracy exceeding 90% with only eight person-days of tuning. We also gave an overview of our automatic FAQ chatbot building system, which is based on the proposed technology.

In the future, we intend to further improve accuracy and performance of the proposed technology and system, based on feedback we obtain through commercial services.

REFERENCES

[1] Repl-AI Web site.
https://repl-ai.jp/

[2] IBM: "Conversation | IBM Watson Developer Cloud."
https://www.ibm.com/watson/jp-ja/developercloud/conversation.html

[3] Microsoft: "Azure Bot Service - Chatbot | Microsoft Azure."
https://azure.microsoft.com/ja-jp/services/bot-service/

[4] Microsoft: "QnA Maker API | Microsoft Azure."
https://azure.microsoft.com/ja-jp/services/cognitive-services/qna-maker/

[5] K. Onishi et al.: "Natural-language Dialogue Platform for Development of Voice-interactive Services," NTT DOCOMO Technical Journal, Vol.17, No.3, pp.4–12, Jan. 2016.

Technology Reports | Demand Prediction | Mobile Kukan Toukei (Demographic Stats from Mobile Phones) | Deep Learning

Special Articles on AI Technologies Contributing to Industry and Society

AI Taxi —Taxi Passenger Demand Prediction Technology for Optimizing Traffic—

Service Innovation Department Satoshi Kawasaki Shin Ishiguro Yusuke Fukazawa

IoT Business Department Masanori Fujita Ryohei Suzuki Akihito Makishima

To increase taxi driver revenue, it is important for them to find customers efficiently, even in unfamiliar areas and time periods, and to maximize the time that the taxi is occupied. To address this, NTT DOCOMO has developed technology to predict demand for taxis by area, based on past taxi operation data and statistical information on locations of people. The technology is being offered commercially as a service. The technology enables taxi drivers to check changes in demand in real time, increasing productivity through efficient operation, and also could help reduce taxi customers wait times.

1. Introduction

To improve revenue in taxi operation, it is very important to reduce the amount of time taxis are vacant, and maximize the time that they are occupied. Taxi drivers can hope to increase their revenue if they are able to find customers efficiently and increase their occupancy rate, even in unfamiliar areas and time periods and in environments that change constantly, with opening of new roadways and commercial facilities, and conditions that change suddenly, such as when a train is delayed. Improvements in efficiency will also improve productivity, and promote a more work-friendly environment.

In response to these conditions, NTT DOCOMO has devised a method for predicting taxi demand in any given area using data on past taxi rides and localized population statistics*.

* The population statistics used by AI Taxi indicate collective numbers of people for a given area or other attribute, and contain no information that can be used to identify an individual. As such, there is no way for anyone to learn anything about anyone else's movements from these population statistics. The population statistics used by AI Taxi conform to the Mobile Kukan Toukei guidelines indicated below. https://www.nttdocomo.co.jp/corporate/disclosure/mobile_spatial_statistics/guideline/index.html

We also developed a demand-prediction system incorporating the method, and conducted practical testing to verify its utility in cooperation with a taxi operator. The tests demonstrated effects of improving revenue and occupancy rates in all time periods of the trial. We have begun offering this system commercially as our AI Taxi®*1 [1] product. By providing demand prediction system results to taxi drivers, taxis can be dispatched to areas where demand is high, as shown in **Figure 1**. This results in reduced wait times for customers, as shown in **Figure 2** (1), and can also lead to increased revenue for the taxi business, as shown in Fig. 2 (2) - (4).

2. Taxi Passenger Demand Prediction Technology

2.1 Overview

The technology takes input data from Mobile

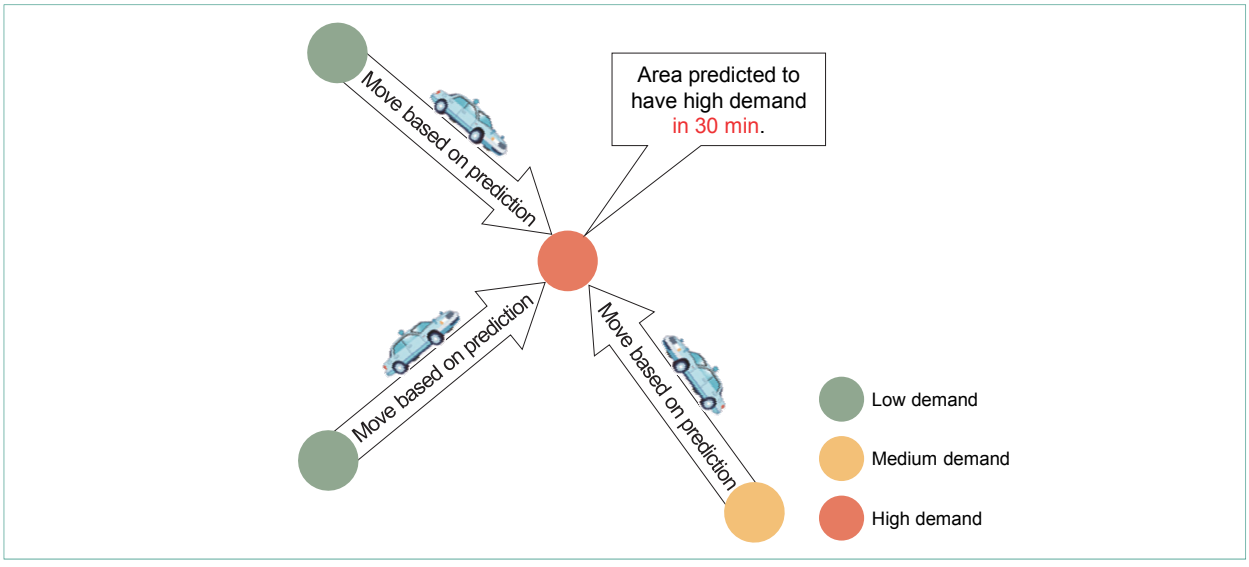


Figure 1 Demand estimation and taxi movement

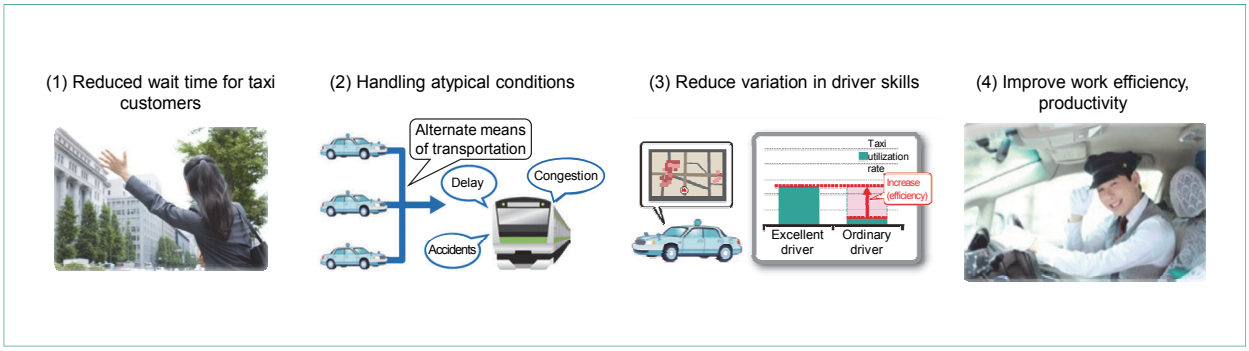


Figure 2 Desired effects from AI Taxi

*1 AI Taxi®: A registered trademark of NTT DOCOMO, Inc.

Kukan Toukei (Demographic Stats from Mobile Phones) Real-time Version*2, weather forecasts, past taxi ride data and other sources, and predicts taxi demand on a 500 m mesh*3 for 30 minutes in the future (demand is the number of rides to be taken within a given section of the mesh).

To predict future time-series data from past records, generally either an Auto Regressive (AR) model*4, or a further development of AR called an Auto Regressive Integrated Moving Average (ARIMA) model*5 is used. These time-series predictive models could also be used for taxi demand, but taxi demand is also affected by other factors, such as weather forecasts and fluctuations in the local population, so we have used a multivariate auto-regressive model*6 that incorporates this additional data to improve accuracy. However, with auto-regressive models, including multivariate auto-regressive models, suitable weightings (parameters) of the feature values for the model equations at each mesh point are generally decided by a person. Expanding as a service over wider areas would increase the number of mesh points, so a method for determining such weightings mechanically was needed. As such, we used deep learning techniques to obtain combinations of data and parameters that can improve accuracy from such heterogeneous data. The commercial service utilizes the result from the multivariate auto-regression model or deep learning, whichever yields the more accurate result, as shown in **Figure 3**.

Deep learning is a machine learning*7 technique that requires much processing power for the learning process, but is attracting attention recently because increases in hardware performance have made it possible to implement the technique at lower cost than before. Deep learning has already demonstrated dramatically better performance than other machine learning methods in various fields such as image recognition, natural language processing,

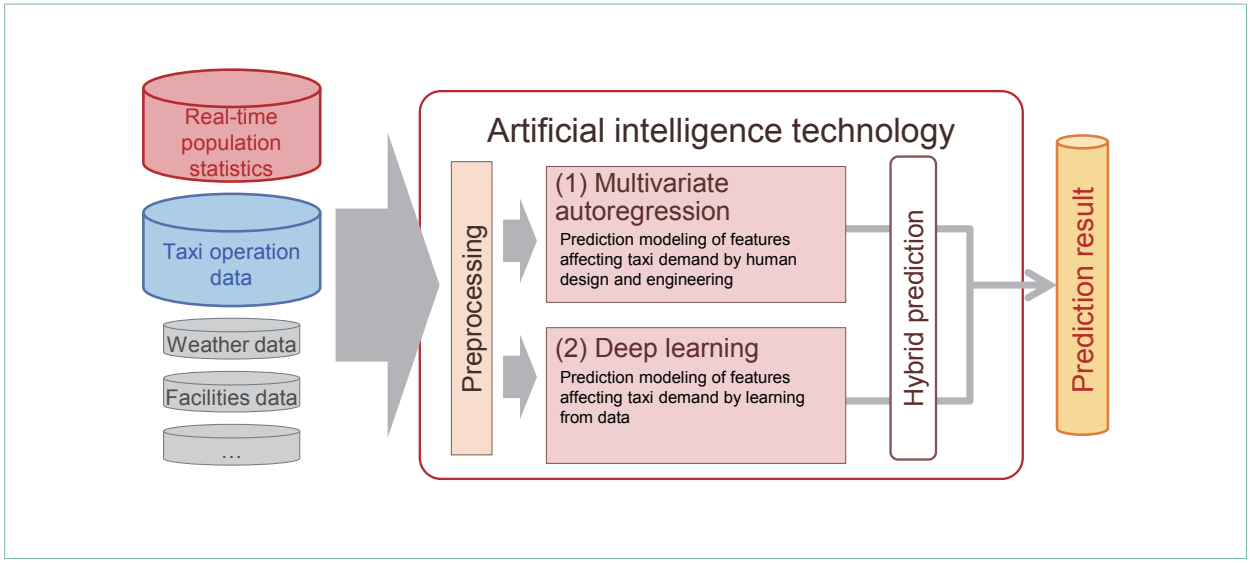


Figure 3 Taxi demand prediction technology architecture

*2 Mobile Kukan Toukei, Real-time Version: Population statistics generated using the NTT DOCOMO mobile telephone network, able to estimate recent population values on a 500 m grid every ten minutes.

*3 Mesh: A grid dividing the country into square sections, along lines of latitude and longitude.

*4 AR model: A regressive model, in which result values (objective variables) are modeled using an equation expressed in terms of affecting values (explanatory variables), but one in which past values of objective variables are used as explanatory variable.

*5 ARIMA model: A regressive model combining Auto Regressive, Moving Average, and Integrated models. The moving average model sums the differences between past predicted values and actual values, and the integrated model uses the differences for modeling.

and speech recognition. Deep learning uses a neural network*8 with many layers (generally four or more) and is able to extract important elements from data, expressing structure and relationships within the data as high-level feature values from very simple feature value inputs, without the need to extract or select advanced feature values or otherwise design a model beforehand.

2.2 Using the Real-time Version of Mobile Kukan Toukei

A major feature of this technology is that it uses the real-time version of Mobile Kukan Toukei to predict demand. Mobile Kukan Toukei*9 has been used earlier with past population statistics [2], but to improve the accuracy of taxi demand predictions it is important to use population statistic closer to the current time, so we have used the real-time version of Mobile Kukan Toukei, which can provide such recent data.

As shown in **Figure 4**, taxi occupancy rates are higher at times and in locations where population is concentrated more than usual, so use of population data could help in finding latent demand.

There are correlations between population fluctuations and fluctuations in demand for taxis, as shown in **Figure 5**. At mesh point A, as the population increases, demand for taxis also increases. It is likely that this mesh point includes other transit facilities, such as a train station, and demand increases (high values in the graph) due to people coming from other areas and transferring to taxis. Conversely, at mesh point B, increases in demand for taxis occur roughly five hours after increases in population. It is likely that there is a commercial establishment or event being held at this mesh point, and visitors to the facility spend some time there, and then take a taxi to some other area. This illustrates how changes in population as input data could be used to predict future changes in demand for taxis. However, depending on the characteristics of each area, the delay between

Figure 4 Taxi customer acquisition rate

*6 Multivariate auto-regressive model: An auto-regressive model extended for multiple variables. Also called a vector auto-regressive model.

*7 Machine learning: A framework that enables a computer to learn useful judgment standards through statistical processing from sample data.

*8 Neural network: A type of model that imitates characteristics of the human brain using computers. They use numerical models that can adjust the degree of connection between nodes using training data, similar to how repeated training changes the strength of connections between neurons in the brain. For machine learning in the past, a designer would have to find characteristics in the input data and design feature values, but a big difference with neural networks is that the training data is given to the neural network, and the feature values are obtained mechanically, similar to what happens in the human brain.

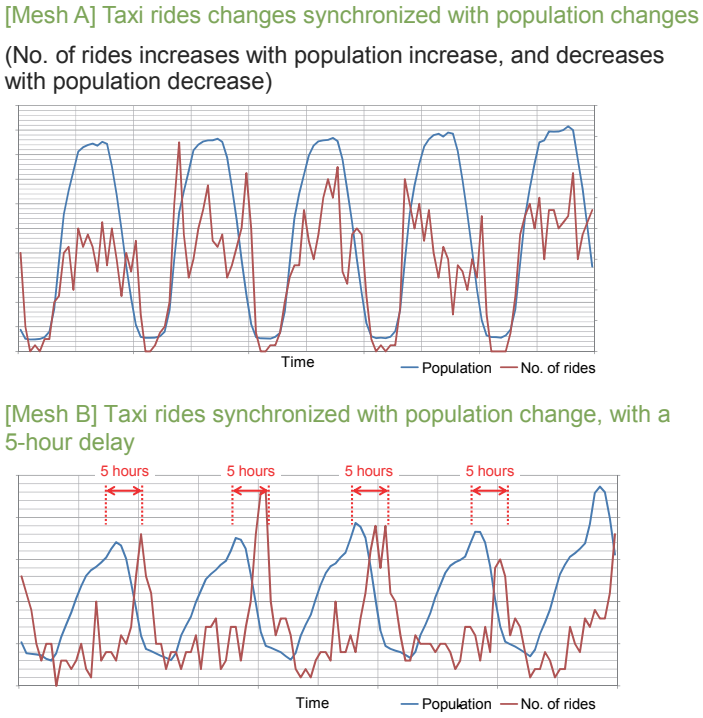


Figure 5 Relationship between changes in taxi demand and population

changes in population and demand for taxis can be different, so to make predictions, we need to determine how much emphasis should be placed at various points earlier in the data for each area. To handle such complex interrelations within the data, we decided to use deep learning and its ability to extract characteristics mechanically.

2.3 Demand Prediction Using Deep Learning

Various network structures have been investigated for deep learning, but we have used the Stacked denoising Autoencoders (SdA) model [3], which develops the autoencoder [4] idea further, in our method.

An autoencoder is a network structure that reduces the input data to intermediate data with fewer elements than the input, and then restores it to the same number of elements as the input. It consists of three layers: the input layer representing the input data, a hidden layer representing the intermediate data, and an output layer representing the restored data. The hidden layer has fewer elements than the input data, so when restoring the data, the weightings for less important elements have been reduced, and the weightings

*9 Mobile Kukan Toukei: A registered trademark of NTT DOCOMO, Inc.

for elements that have particular importance in reproducing the input data get increased. The network structure used for our method stacks multiple autoencoders, as shown on the right side of **Figure 6**, enabling it to extract feature values more efficiently and achieve higher accuracy. Important features are extracted through several hidden layers. Also, to handle demand prediction as a regression problem*10, the last autoencoder hidden layer, which extracts the most important feature values, is connected to an output regressive prediction layer*11. In this layer, the values are multiplied by the learned weightings, and the results are added together. Then the final demand prediction results are output. In addition to past taxi ride and population time-series data, statistical data, such as average rides for the same day of the week and time of the day, is also added to the input layer, as shown on the left in Fig. 6. Note that in the training process with an SdA model, relations between temporal and spatial information are learned mechanically from large amounts of training data, in the form of the weightings on links between each node.

Adding noise to the autoencoder input during training also contributes noise-cancelling (denoising) characteristics. The autoencoders are intended to reconstruct the input data, but by adding faults (or noise) to the input data, and training them to restore the original data which has no faults (or noise), we can expect the effects of faults or noise in the original data to be minimized, and the information

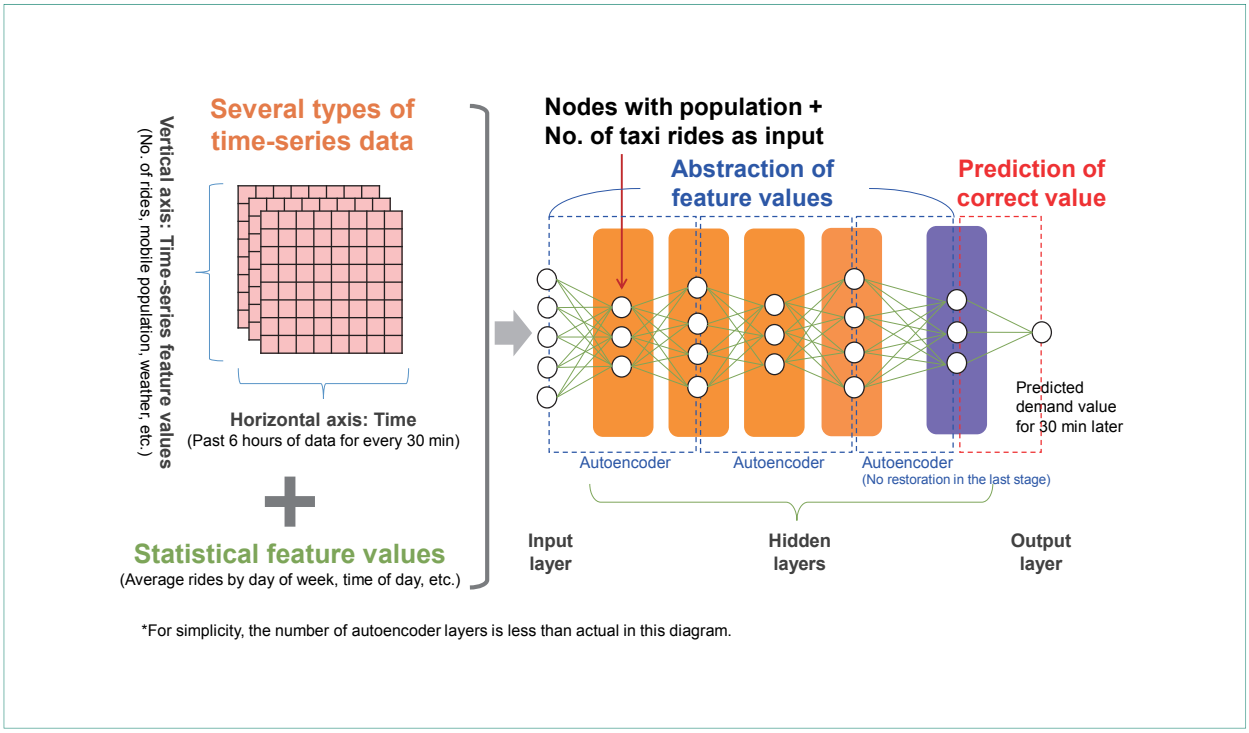


Figure 6 Network structure for predicting taxi demand

*10 Regression problem: A problem that can be handled as a regression, deriving the objective variable using with an expression in terms of the affecting variables.

*11 Prediction Layer: A layer that estimates what will appear next from a given input.

most important for recovering the original data to be extracted preferentially.

3. Testing

We conducted tests predicting taxi demand in Tokyo using this technology. We computed how many taxis were expected to find customers in each area on a 500 m mesh in the following 30 minutes. We used data from April 1, 2015 to August 31, 2016 as training data, and from September 1 to September 14, 2016 as evaluation data. Other details regarding the data are given in **Table 1**. Tablets next to the driver's seat displayed a screen like that shown in **Figure 7**, giving the taxi demand prediction results for areas on a map delineated by red lines and enabling drivers to move to areas where predicted demand was higher.

During the testing period, there was some variation in accuracy of predictions by the technology depending on area, but results ranged from 93 to 95%, confirming that the predictions were very accurate. The actual occupancy rates and predicted results are compared in **Figure 8**. Compared with

Table 1 Data used in verification testing

Data source		Metropolitan Tokyo
Time period		Apr. 1, 2015 – Sep. 14, 2016
Taxi data	No. of cars	4,400
	Acquisition frequency	Every 5 to 10 s
Population data	Resolution	500 m
	Acquisition frequency	Every 10 min
Rainfall data	Resolution	250 m
	Acquisition frequency	Every 10 min

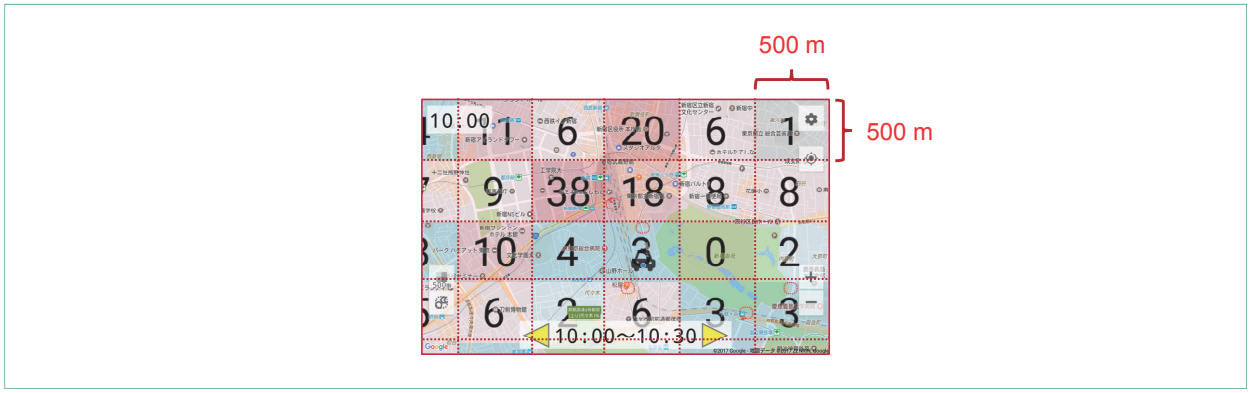


Figure 7 Screen displaying demand prediction results

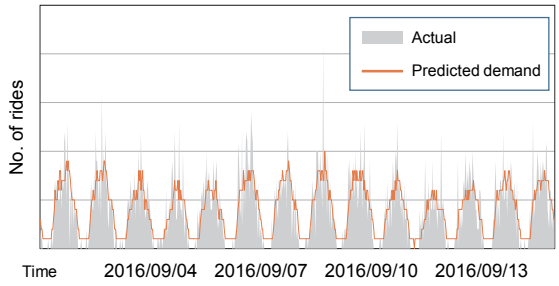


Figure 8 Comparison of demand prediction results and actual rides

results one month before the tests, sales amounts for drivers using the demand prediction system increased by an average of 1,400 yen per driver per day in all times tested, which was more than the average increase for all drivers. The taxi occupancy rates, which are an index of efficiency in getting customers, also increased.

4. Conclusion

This article described a technology used in the AI Taxi Service that predicts future demand for taxis. Matching supply with demand is important for achieving efficient operation. The technology enables the supply of unoccupied taxis to be optimized by predicting demand ahead of time. In future work, we will also study dispatch control mechanisms that consider global optimization.

REFERENCES

[1] AI Taxi Service Web site: https://www.nttdocomo.co.jp/biz/service/aitaxi/

[2] Okajima et al.: ""Mobile Spatial Statistics" Supporting Development of Society and Industry - Population Estimation Technology Using Mobile Network Statistical Data and Applications -," NTT DOCOMO Technical Journal, Vol.14, No.3, pp.4–50, Jan. 2013.

[3] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio and P. A. Manzagol: "Stacked Denoising Autoencoders: Learning Useful Representations in a Deep Network with a Local Denoising Criterion," Journal of Machine Learning Research, pp.3371–3408, Dec. 2010.

[4] G. E. Hinton and R. R. Salakhutdinov: "Reducing the dimensionality of data with neural networks," Science, Vol.313, Issue 5786, pp.504–507, 2006.

Technology Reports
Image Recognition Deep Learning AI
Special Articles on AI Technologies Contributing to Industry and Society

A Retail Shelving Analysis Solution Using Image Recognition —Recognizes Shelving Allocation and Quantifies Inventory by Analyzing Photos of Shelved Merchandise—

Service Innovation Department Hayato Akatsuka† Issei Nakamura
Kansai Branch, Corporate Sales Dept. Sungmyeong Koh

In consumer retail, it is important to understand the condition of product display (planograms, retail space management) when analyzing factors affecting sales. Conventionally, obtaining planogram data has been done manually, but it has been costly and involved a heavy workload for employees. NTT DOCOMO has developed retail shelving image recognition technology and used it to implement creation of planogram data automatically, from photographs of retail shelving taken with a smartphone or other camera. This greatly reduces the amount of time required for this work. A solution using this technology was awarded the 19th Automatic Identification Systems Prize, Award of Excellence, from the Japan Automatic Identification Systems Association in 2017, and has been provided by NTT DOCOMO to partner enterprises since April 2018.

1. Introduction

In consumer retail, it is important to know the condition of product displays (planograms[*1]) when analyzing factors affecting sales. Currently, consumer goods manufacturers create planogram data manually, writing documents from notes taken by sales representatives[*2]. With the recent growth in large-scale retailers, sales representatives need to visit an increasing number of stores, and with the continually shrinking workforce, a shortage of workers can be expected in the near future. To solve

† Currently Innovation Management Department

*1 Planogram (tana-wari): Refers to the layout of how products are displayed on retail shelving.

this problem, it is critical to reduce operating costs and improve work efficiency. We have estimated that there are tens of thousands of sales representatives performing such tasks in Japan, and we can see huge demand in the market for automating creation of planogram data.

We have proposed a method using image recognition technology to automatically extract planogram data from photographs of retail shelving. There were two main issues with using existing image recognition technologies in real retail environments.

- The first is that object detection is difficult for products when they are packed tightly onto compact displays*3. When shelving space is limited, as it often is in Japan, products like shampoo refill pouches are placed in compact displays and possibly only the part of product face is shown or the shape of the product is distorted.
- The second is that it may be difficult to uniquely identify products because they are not always facing the front. In a real store, a product could be facing in any direction on the shelf.

In earlier solutions, the photographing process was manually adjusted in order to make image recognition easier. Product overlap was eliminated and packages were faced to the front as much as possible. However, such adjustment was time consuming and not practical.

For these reasons, NTT DOCOMO has developed a retail shelving image recognition engine that resolves issues with compact displays and photographing angle, using deep learning technology to analyze photographs of retail shelves automatically and create planogram data. Our image recognition engine was recognized by the Japan Automatic Identification Systems Association for "Innovative reproduction of retail shelving using image recognition" and "Solution feasibility," receiving the 19th Automatic Identification Systems Prize, Award of Excellence in 2017 [1]. It has been provided to Cyber Links Co., Ltd. since April 2018 as its first user [2]. This article describes technical details of our image recognition engine together with practical use cases.

2. Image Recognition Overview

Our image recognition engine is composed of two technologies.

The first is a deep-learning based object detection technology that detects whether shelves and products are present in a photograph and determines the spatial position of each product region. The outlines shown in **Figure 1** are the results of object detection. NTT DOCOMO's object recognition technology uses deep learning trained with images of retail shelves in various states, so products can be detected accurately, even if they are packed into shelves in small spaces.

The second is a specific-object recognition technology that uses local feature values*4 and is able to identify the product in each object region on the shelf. Product region partial images detected as described above are compared with a large number of product images stored in an image database. Images taken from various angles are stored in a large image database beforehand, which enables

*2 Sales representative: An employee that travels to supermarkets, department stores and other retail outlets to promote sales of a company's products. Specific duties could include making sales recommendations and understanding display conditions of the company's products in the stores, studying products from other companies, and getting feedback and any requests from the stores.

*3 Compact displays: A way of completely filling shelves in retails spaces with limited space, often resulting in products becoming somewhat compressed.

*4 Local feature values: Extracted from data, values (numbers) that characterize the data. In this article, "feature values" refers specifically to image feature values, which are characteristic points (corners) extracted from the image and the surrounding distribution of brightness.

Figure 1 Retail shelving image recognition process

our specific-object recognition technology to recognize products accurately, even if the image is not taken from the front.

The processes used by these technologies are shown in Fig. 1.

2.1 Object Detection Technology

1) Algorithm Details

The object detection technology developed by NTT DOCOMO is an application of deep learning and is able to estimate the position (coordinates of upper left and lower right of a rectangle) and a category for each object in an image. To do so, the object detection engine must be trained beforehand in a machine learning*5 process for the category of objects to be detected. For this process, we prepared hundreds to thousands of images with annotation data*6. The annotation data consisted of position data, indicating where the object is depicted in the image (upper-left (x_{min}, y_{min}) and lower-right (x_{max}, y_{max}) coordinates of rectangle) and the object category (e.g.: fabric softener, laundry detergent, drinking water, shelf, etc.) as shown in **Figure 2**.

The machine learning process creates a trained model, which is loaded into the object detection engine, which then performs its inference*7 process to detect objects in input images.

The inference process consists of the following four phases (**Figure 3**).

(1) Extract feature values

Input images are converted to feature values through several convolution*8 layers and pooling*9 layers.

(2) Estimate candidate regions

Using the feature values extracted above, multiple regions are cut out at various aspect ratios and scales, and a fully-connected layer*10 is applied to each to predict whether it is object or background.

*5 Machine learning: Technology that enables computers to acquire knowledge, decision criteria, behavior, etc. from data, in ways similar to how humans acquire these things from perception and experience.

*6 Annotation data: In this article, refers to metadata indicating what is in an image.

*7 Inference: In this article, refers to use of a previously trained model to predict what is depicted in an image.

*8 Convolution: A process of scanning an input such as image or feature value horizontally and vertically, multiplying by a vector of certain size and outputting the value. Extracts patterns similar to the vector used.

Annotation

Object	Type	Position data			
		x_{min}	y_{min}	x_{max}	y_{max}
1	Fabric softener	640	30	700	460
2	Laundry soap	…	…	…	…
3	Laundry soap	…	…	…	…
…	…	…	…	…	…

Figure 2 Example of image and annotation needed for training

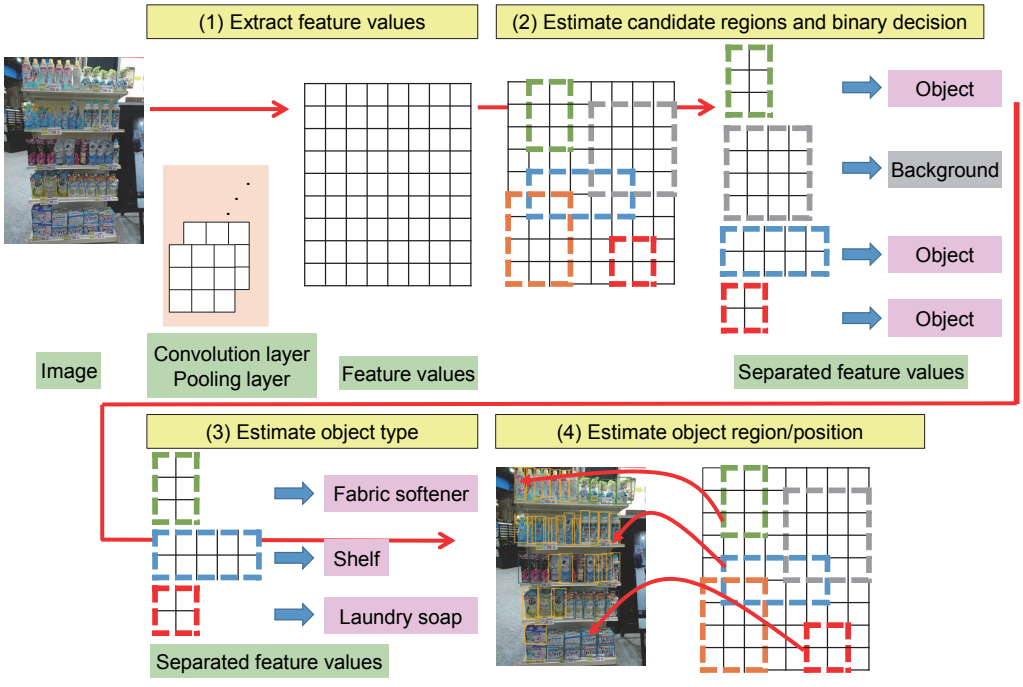


Figure 3 Detection algorithm process

*9 Pooling: A process of scanning against a feature value horizontally and vertically, and outputting the maximum or average value within a fixed size (2 × 2, 3 × 3, etc.). Reduces dimension of the feature value and increases robustness of inference.

*10 Fully-connected layer: Weighting in which all feature values are multiplied by a weighting and added to output a single value. Feature values extracted using convolution are converted to product categories by multiplying them by these weightings.

(3) Estimate object category

For regions judged to be "object" in (2), multiple fully-connected layers with different weights than in (2) are applied to infer the object category. This yields multiple candidate categories and corresponding probabilities, and the region is assigned the category with the highest probability.

(4) Estimate location of object region within the image

The position of the object within the image is estimated using the object category probability found in (3), the position in the image, and the feature values.

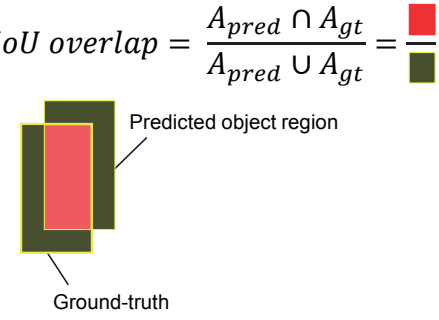
The above process is able to estimate the position and the category of multiple objects depicted in a single image.

Note that the method is able to detect shelves as well as products. This enables it to also estimate factors such as on which shelf a product is placed, and how high products are stacked.

2) Recognition Accuracy

We performed the machine learning process using several thousand images of retail shelves with many products on them, and then evaluated the accuracy of object detection by applying the object detection engine to 100 images taken in real stores and not used for training. Products were in two categories: laundry products such as laundry detergent and fabric softener, and beverages such as drinking water in PET bottles and cans of beer. Recall (No. of products correctly detected/total no. of products) and Precision (No. of correctly detected products/No. of detected products) were used as indices of accuracy. To evaluate accuracy, the corresponding ground-truth region (correct region) for each detected region must first be selected. To do so, an index called IoU overlap is used, which expresses the amount of overlap between two regions as a value from 0.0 to 1.0, as shown in **Figure 4**. The higher the value of IoU overlap, the greater the amount of overlap between the two regions. Detection is judged to be

$$IoU\ overlap = \frac{A_{pred} \cap A_{gt}}{A_{pred} \cup A_{gt}} =$$



A_{gt} : *Area of ground-truth region*
A_{pred} : *Area of predicted object region*

Figure 4 IoU overlap

correct only if the IoU overlap is 0.6 or greater, and the predicted object category is the same as the object category of the ground truth. However, if the probability of the object category computed in (3) was below 0.6, the confidence level is considered low, and such cases are excluded.

When evaluating under these conditions, detection accuracy for beverages had recall of 91.2% and precision of 92.1%. Laundry product detection accuracy had recall of 92.0% and Precision of 99.7%. As shown in **Figure 5**, product packages on real store shelves were not all facing the front, products in pouch-type packaging were easily deformed, and packages of the same product overlapped and could be partially hidden, but the technology was still able to detect them accurately.

3) Processing Speed

We measured processing speed for detection using a GPU and using a CPU. Processing completed in approximately 0.3 s when performing inference with a single NVIDIA Tesla*11 M40 GPU. On the other hand, performing all computations on the CPU (Intel® Xeon®*12 CPU E5-2630L v3 @ 1.80 GHz) required approximately 7.0 s, showing that speed increases greater than ten-times can be achieved using a GPU. System requirements differ according to the application scenario, but this suggests that with the current system, use of a GPU will be essential for real-time processing. Note that the processing speeds described here depend on the size of the input image files, so shorter processing times may be possible using low resolution images.

2.2 Specific-object Recognition Technology

1) Algorithm Details

The specific object recognition technology compares input images with images pre-registered in

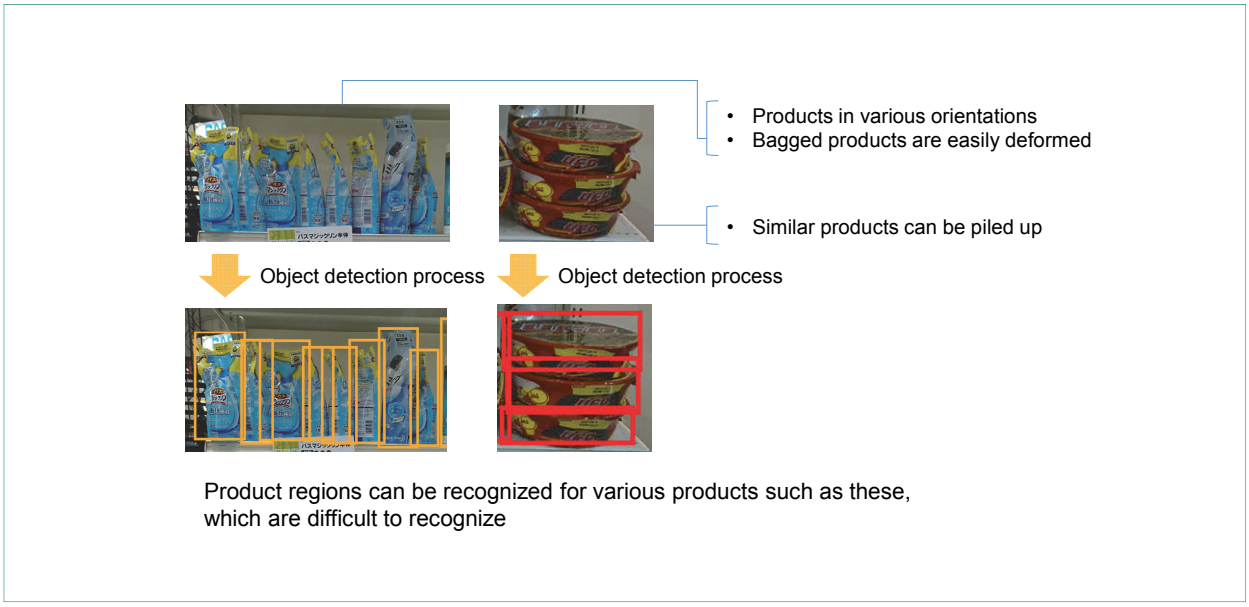


Figure 5 Features of products on retail shelves

*11 NVIDIA Tesla: A trademark or registered trademark of NVIDIA Corporation in the USA and other countries.

*12 Intel® Xeon®: A trademark of Intel Corporation or a subsidiary in the USA and other countries.

an image database, and finds a registered image that is similar to the input image (**Figure 6**).

When recognizing the retail shelf image in Fig. 1, the partial images in the regions found by object detection are input to specific-object recognition engine. The engine is able to recognize products displayed at various angles by comparing with pre-registered images of the products taken from various angles in the database. For details of NTT DOCOMO's specific object recognition algorithm, see reference [3].

2) System Requirements

NTT DOCOMO's specific-object recognition technology is able to process images in real time, even with several million images registered in the image database, but to do so, the database must be loaded in memory. The amount of memory required depends on the number of images and the data-size allocated to each image, but there is an upper limit to the amount of physical memory that a single server can have, so there is a limitation on the number of images a server can handle. As such, a large-scale image database can be built by scaling out*13 with multiple servers. Specific-object recognition can operate at high speed on a CPU, recognizing an image in a few hundred milliseconds, depending on engine settings and the number of registered images.

3) Recognition Accuracy

We evaluated recognition accuracy by preparing retail shelves in a test environment reproducing scenes of products displayed in a real store. Types of products we used included noodles (cup noodles, bagged noodles), detergent (bottles, refill pouches), and beverages (cans, PET bottles). In results of evaluating several hundred product images, the top-ranked candidate was correct 95.96% of the time, demonstrating very high accuracy.

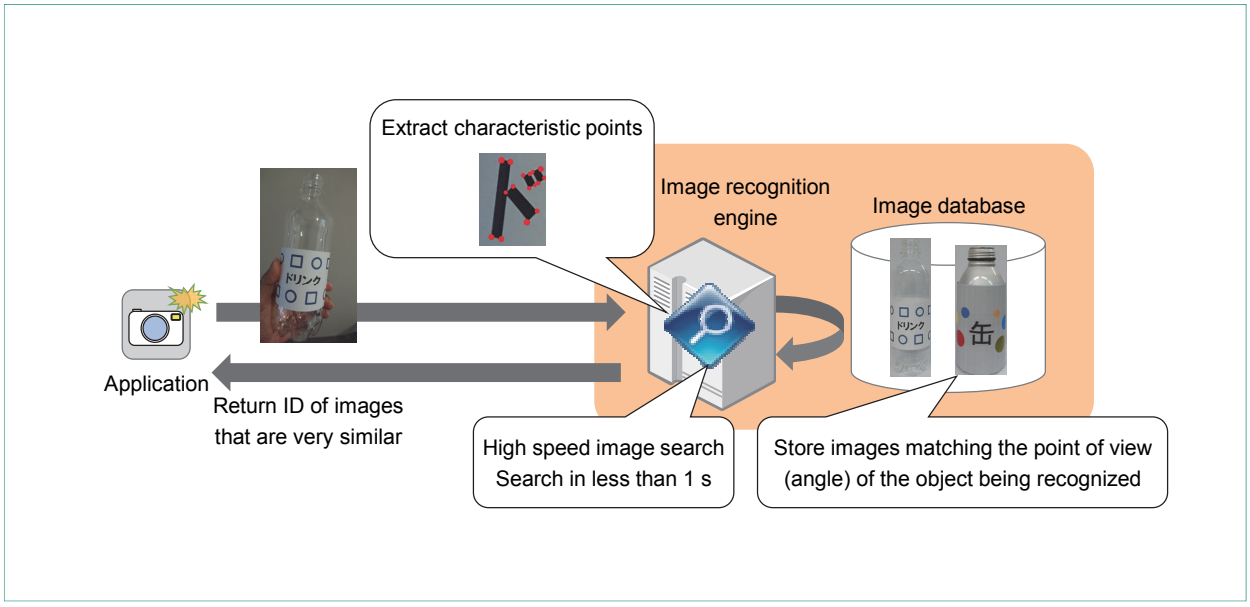


Figure 6 Specific object recognition

*13 Scale out: Adding and assigning new resources to increase processing capacity when service requests increase and there is insufficient processing capacity on the network.

On the other hand, recognition failed in cases where the size or color of the product was wrong, or the image was poor due to camera shake. Products of different size or color, or where just one character is different, are so similar that a person could also easily make a mistake in distinguishing them, so they are generally difficult to distinguish with image recognition as well. For camera-shake, the necessary information cannot be extracted from the image, and this degrades recognition accuracy. Besides these cases, recognition accuracy tended to drop when the image resolution was low, and when the product region in the image was small, so that very little image information could be extracted. To increase recognition accuracy, camera shake must be minimized, and retail shelves must be photographed with as high resolution as possible.

3. Creating Planogram Data

The image recognition engine can be used to identify products in photographs of retail shelving, to analyze the state of product displays, and to output the results as planogram data. This planogram data can be loaded into planogram simulation software to visualize and analyze planograms (**Figure 7**). Typical planogram software used in Japan includes Tana POWER®*14 [4] from Cyber Links Co., Ltd. and StoreManager®*15 [5] from Nippon

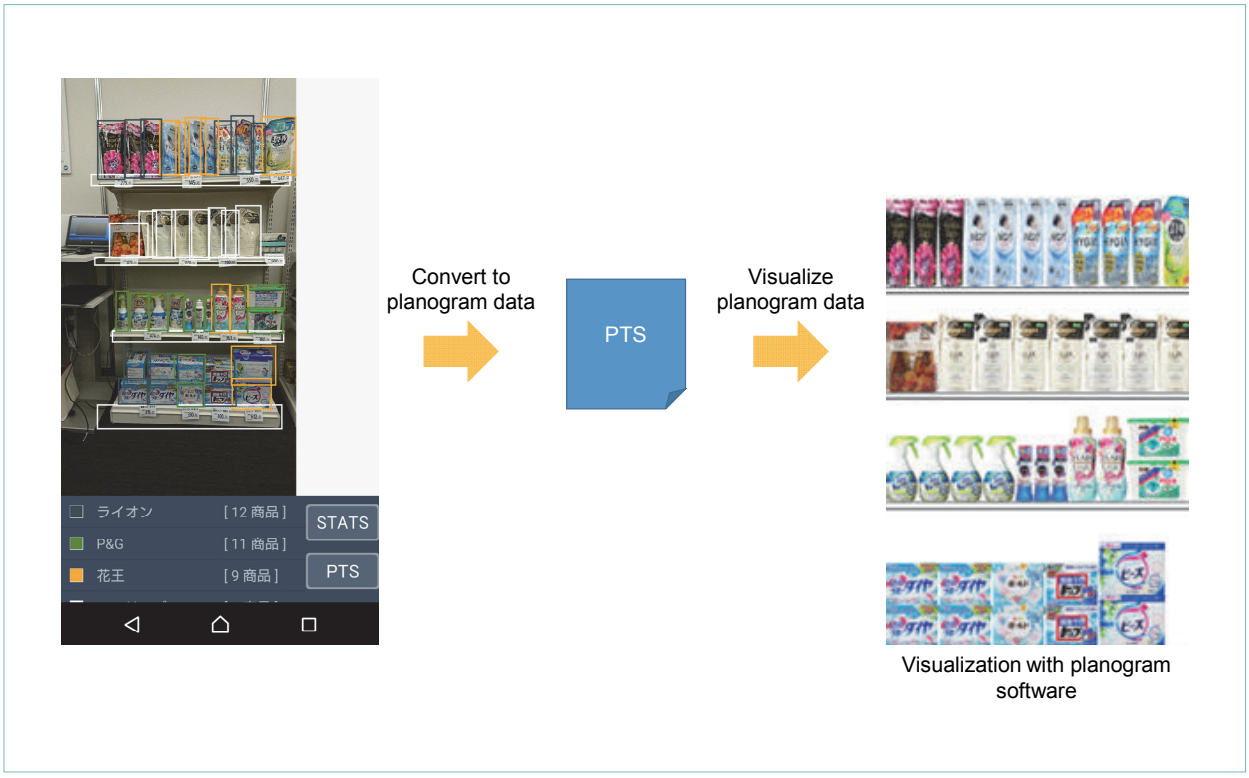


Figure 7 Visualization of planogram data

*14 Tana POWER®: A registered trademark of Cyber Links Co., Ltd.
*15 StoreManager®: A registered trademark of Nippon Sogo Systems, Inc.

Sogo Systems, Inc.

In Japan, a common format used for representing planogram data is Planogram Transfer Specification (PTS) [6]. PTS includes information needed for visualizing products, such as the number of faces*16, stacked height, and which shelf the product is on for each product. The object detection technology described above is able to detect shelves as well as products, so it is able to count number of faces and stacked height accurately for each product on each shelf, and to create planogram data conforming to PTS specifications. This enables the original product displays to be reproduced accurately without needing to reconfigure the shelves when loading data into planogram software and visualizing the planogram.

4. Use Cases

As mentioned earlier, this image recognition engine can be used to create planogram data for sales representatives. We estimate that by creating planogram data automatically from photographs, work time is reduced by a factor of ten compared to creating it by hand. In addition to reducing labor costs, it can also mitigate heavy labor. Specifically, current methods for creating planogram data involve using a barcode reader, which places a heavy burden on the back and legs due to movement up and down from the top to bottom shelves. Using the image recognition engine, this work can be completed with just two or three photographs from the front, simplifying and reducing the workload.

Other scenarios could include use by sales representatives for other consumer products, or for photographs used by retail staff to create work reports. Currently, shops use photographs of retail shelving to report completion of work preparing product displays, but these are limited to a simple visual check of the photographs. Automatically creating planogram data from retail shelving photographs could help head office or other remote offices with management work, inspecting displays at remote stores, checking for errors, and understanding any new display techniques being used at the stores.

5. Conclusion

This article has described two elemental technologies of a retail shelving image recognition engine: one that detects products on shelves and another that recognizes specific products. It also described a capability to create planogram data automatically, by identifying products from images of retail shelving, and detecting their position and shelf within the display using these technologies.

Further work on this image recognition engine includes testing the technology in real shops and improving accuracy. In particular, further study is needed on specific-object recognition technology that can robustly handle products that are difficult, as described above: differing in size, color, or text. Some stores also have price rails*17 attached to shelves, which hide parts of products and obscure information, degrading recognition accuracy. In the future, we intend to collaborate with more partners, testing this technology and identifying cases that make recognition difficult such as these, and continuing to improve our recognition algorithms.

*16 Number of faces: Distribution term indicating the number of items that can be seen displayed when facing the retail shelf.

*17 Price rail: A rail on each shelf in a product display with price tags attached.

Besides product display image recognition, NTT DOCOMO is also providing image recognition solutions to government and partner enterprises in a wide range of fields such as sports video analysis [7], AR service applications [8], business optimization applications by digitizing name cards [9], and detecting pine wilt in coastal forest reserves using drones [10]. Although there are restrictions, some of the image recognition functionality developed by NTT DOCOMO is published by docomo Developer support [11], as Application Program Interfaces (API)*18 for use in developing applications and services. These can be accessed by simply joining docomo Developer support and applying for access. NTT DOCOMO will continue development on image recognition technologies to provide value to partner enterprises in various fields, including retail shelving image recognition.

REFERENCES

[1] Japan Automatic Identification Systems Association: "Japan Automatic Identification Systems Prize | JAISA Japan Automatic Identification Systems Association." http://www.jaisa.jp/award.php

[2] NTT DOCOMO Press Release: "DOCOMO Launches AI Engine for Fast, Accurate Shelf Analysis —Recognizes shelf allocation by analyzing photos of shelved merchandise—," Mar. 2018. https://www.nttdocomo.co.jp/english/info/media_center/pr/2018/0316_00.html

[3] H. Akatsuka et al.: "High-speed, Large-scale Image Recognition API," NTT DOCOMO Technical Journal, Vol.17, No.1, pp.10–17, Jul. 2015.

[4] Cyber Links Co., Ltd: "Basic Operation | Tana POWER | Cyber Links Co., Ltd. Tana POWER/Mise Power." https://www.tanapower.com/tpower/basic_operation.htm

[5] Nippon Sogo Systems, Inc.: "StoreManager | Nippon Sogo Systems Inc." https://tanawari.jp

[6] Japan Planogram Research Association: "Introducing Planogram Transfer Specifications (PTS): Japan Planogram Research Association." https://www.planet-van.co.jp/planogram/pts/index.html

[7] Soccer.com, NTT DOCOMO: "AI Technology "Sports Video Sensing" Development, Multi-angle Automatic Video System linked trials begin —Providing a new sports experience enabling review of one's own plays easily on video—," Feb. 2018. https://www.nttdocomo.co.jp/binary/pdf/info/news_release/topics/topics_180201_00.pdf

[8] NTT DOCOMO: "Cybernet Systems provides image recognition system to cybARnet," Oct. 2015. https://www.nttdocomo.co.jp/binary/pdf/corporate/technology/rd/topics/2015/topics_151510.pdf

[9] NTT DOCOMO Press Release: "Sansan Inc. adopts NTT DOCOMO image recognition system," May 2016. https://www.nttdocomo.co.jp/info/news_release/notice/2016/05/25_00.html

[10] NTT Press Release: "Collaboration agreement reached on drone test project in Niigata City," Sep. 2016. https://www.nttdocomo.co.jp/info/news_release/2016/09/21_00.html

[11] NTT DOCOMO: "Image Recognition | docomo Developer support | NTT DOCOMO." https://dev.smt.docomo.ne.jp/?p=docs.api.page&api_docs_id=102

*18 API: An interface that enables software functions to be used by another program.

Technology Reports
Communication Devices | Open Innovation | Design Ideas

"petoco": A Home Communication Device

Communication Device Development Department Keiichi Murakami

At NTT DOCOMO, we have been studying communication among families. As a result, we have identified problems in dual income families, where a child sometimes returns home while both parents are still out, and has fewer opportunities to talk with his or her parents. We developed petoco®*1 to provide these families with a means of communication that addresses the sorts of situations that can arise from the disjointed life rhythms of different family members. This is an innovative device that embodies the application and development of advanced technology based on concepts created using design thinking methods to solve issues from a human-centered design viewpoint.

1. Introduction

Based on our philosophy of creating a new world of communications culture [1], our aim at NTT DOCOMO is to resolve communication-related issues from the user's point of view. In line with this philosophy, we are focusing on the rapid changes that are taking place in family life, especially in Japan [2], and with families as our target users, we have begun a study of new services to address issues in family communication.

For this study, we have taken a new approach. Until now, many businesses have taken new technologies and market trends as their starting point, but it has been difficult to come up with new concepts that go beyond existing products and services, many of which have turned into standardized commodities [3]. In response to this situation, we have been considering more human-oriented approaches in order to identify new latent needs

*1 petoco®: A registered trademark of NTT DOCOMO, Inc.

that users themselves may be unaware of, and to discover new markets based on these needs [4]. One such approach is that of design thinking [5]. As shown in **Figure 1**, the design thinking approach consists of a series of flows between the phases of "Inspiration" (careful observing and understanding the user), "Ideation" (creating ideas based on this inspiration), and "Rapid prototyping and verification" (where these ideas are quickly put into practice) [6].

In this article, we describe the sort of processes we considered for petoco as a concept device created through the use of design thinking to solve family communication issues [7], and we describe the technology that was used to implement it.

2. Application of Design Thinking to Concept Creation

Hereinafter, we describe how we applied the design thinking method to the theme of family communication, and how we arrived at the final concept device through each phase of this process—(a) Inspiration, (b) Ideation and (c) Rapid prototyping and verification.

2.1 Inspiration

In creating this concept, we needed to recognize what sort of problems affect communication within families. To this end, we first sought to gain inspiration from family members who are the intended users of this system. In design thinking, "Inspiration" refers to the close observation and understanding of the user. To achieve this, we actually visited the homes of families with a range of different attributes, and we interviewed them regarding their use of communication. In particular, we focused on the communication issues of families categorized as Double Employed With KidS (DEWKS)—i.e., families with children where both

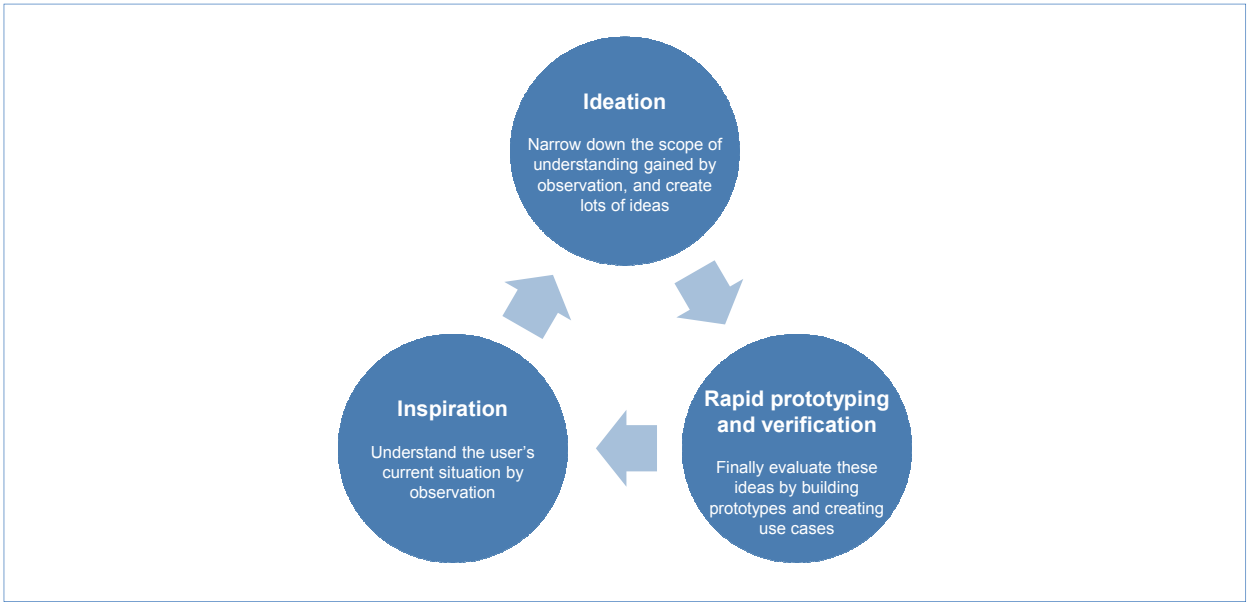


Figure 1 The design thinking process

parents go out to work. Recent surveys have shown that dual-income households are steadily increasing [8], and that further increases in DEWKS are expected in the future. From our interviews we found that DEWKS tend to exhibit characteristics such as the following:

- The lifestyle rhythms of parents and children tend to be highly independent (**Figure 2**)
- Parent-child contact time is very limited
- Since time is limited, they have a very strong desire to communicate with one another
- The lack of time gradually makes communication more troublesome and rarefied

Regarding these problems as a communication gap between parents and their children, we decided to study solutions for closing this gap. We also found that the above problems are more pronounced in DEWKS families with children of primary school age or younger who do not have any means of communication. We therefore conducted further interviews with these families as our target persona[*2]. By interviewing these families at home, we were able to gather specific details on their lifestyle patterns such as the layout of their homes, the arrangement of their furniture and home appliances, and the paths people follow when moving around. From this, we were able to extract true insights[*3] into the subconscious minds of users. As a result, we obtained the interview results and insights shown in **Table 1**.

With reference to these results, we defined our design principles by working towards "Ideation" based on our insights shown in **Table 2**.

Based on these insights obtained in the "Inspiration" phase, we studied device concepts in the

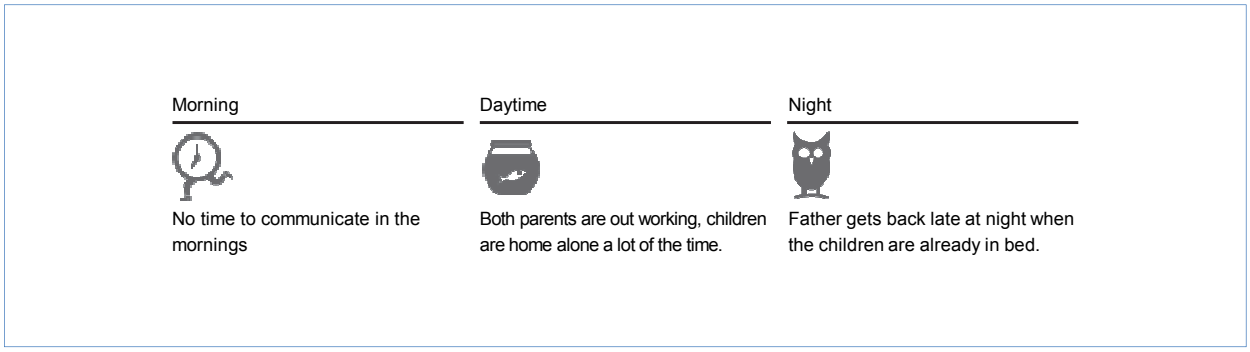


Figure 2 Typical communication gaps in a DEWKS household

Table 1 Interview results and the insights we obtained

Interview results	Resulting insights
While parents continue to expect the same quantity and content of communication from their children, the children are less inclined to talk with their parents (especially their fathers).	Family members find it difficult to talk because their life rhythms are disjointed and they spend little time together. However, they want to talk more about their real feelings.
The use of media such as sticky notes can convey thoughts and feelings that are difficult to express in digital communication.	Even though communication is becoming more rarefied, people can still feel warmth and a sense of connection when communicating through media such as sticky notes and letters.

*2 Persona: An imaginary stereotypical user of products and services.
*3 Insight: The user's intrinsic desires and intentions.

"Ideation" phase.

2.2 Ideation

In the "Ideation" phase, we held workshops and discussions with reference to the design principles extracted in the abovementioned "Inspiration" phase. This process generated a lot of ideas, which we organized and filtered down to refine the device concept through a combination of divergence and convergence. Based on these results, we created the original form of petoco with the product design described below. This included the major themes of a "tangible*4" [9] experience and the role of a "hub" in the design principles obtained from the "Inspiration" phase. We also created use cases for this design.

1) Product Design

Sketches of proposed product designs are shown in **Figure 3**. The features of these designs are presented below.

- An object in the living room or on an important pathway through which families pass
- A utilitarian design that is less susceptible to user preference
- Interactive character attributes with feelings and personality that encourage emotional attachment, particularly by children

2) Use Cases

Some examples of the use cases we came up with are described below.

Table 2 Design principles aimed at gathering insights and developing them into "Ideation"

Insights	Design principles
Use devices to mediate and promote communication within families. Make the product appealing to encourage its use by children.	Provide a hub that can mediate communication between family members whose time is limited.
Design a tangible experience based on existing communication metaphors.	Provide a tangible experience of existing modes of communication such as sticky notes and letters.

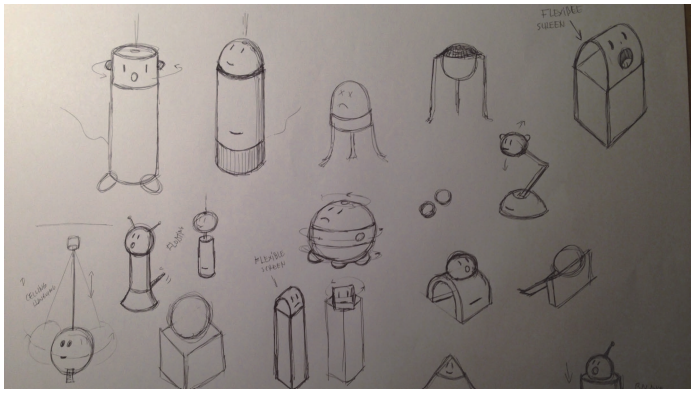


Figure 3 Product design proposal sketches

*4 Tangible: A word used in connection with real objects that can actually be touched.

- When a parent is out somewhere and sends a message home, the message is printed out on a device in the home
- A smartphone can be used to monitor the interior of the home while out somewhere else
- When a child arrives home, it greets them and sends a notification to the parent’s smartphone

2.3 Rapid Prototyping and Verification

At the “Rapid prototyping and verification” phase, we produced several simple mockups[*5] with real textures and a number of functions and gimmicks. By using these mock-ups in the use cases obtained at the “Ideation” phase, we performed a more detailed user survey. Through this survey, we further refined and narrowed down the product design and use cases towards our final concept.

1) Product Design

The evaluation was performed using a mock-up in which a projector was used to create an image of a character on outside of the housing (**Photo 1**). We received a lot of positive feedback regarding the use of a curved product surface on which a character suddenly appears, particularly from children. The users also reported feeling a sense of unity and attachment with the product due to its perceived cuteness, its changes in facial expressions, its responsive behavior, and other such attributes.

2) Use Cases

We decided on the functions of the final design concept after interviewing the users involved in a number of different use cases. Of these, the most positive response was obtained for the ability of this device to pass on messages sent by parents from outside the home. Working mothers were also highly appreciative of the function for monitoring the interior of the home. It was also pointed out that the ability to perform video calls was an attractive additional feature. We also found that ease of use and maintenance-free operation were also important points for prolonged use.

2.4 Final Concept

Through the “Rapid prototyping and verification” phase, we came up with petoco as our final concept. In this concept, we included the basic functions of text, voice and video messaging and video chat between petoco and family members with smartphones who are away from home, and between family members at home using petoco to pass on messages (**Figure 4**). In addition, the product design was made to blend in with the home interior, and attachment was stimulated by

Photo 1 The product undergoing evaluation

*5 Mockup: A model created for the purpose of prototyping and evaluating the appearance of a design.

projecting a character onto the outside of the product (**Photo 2**).

3. Functions Implemented by petoco

The final concept must be embodied as a commercial product and service using feasible technology. This section describes how we implemented the characteristic experience of petoco. The main specifications are shown in **Table 3**.

3.1 Video Output to a Smooth Curved Surface

In the evaluation mock-ups, we simply used an external projector to display images on the outside of the product, but for the final concept device, these images had to be produced completely by the product itself. We therefore adopted a rear

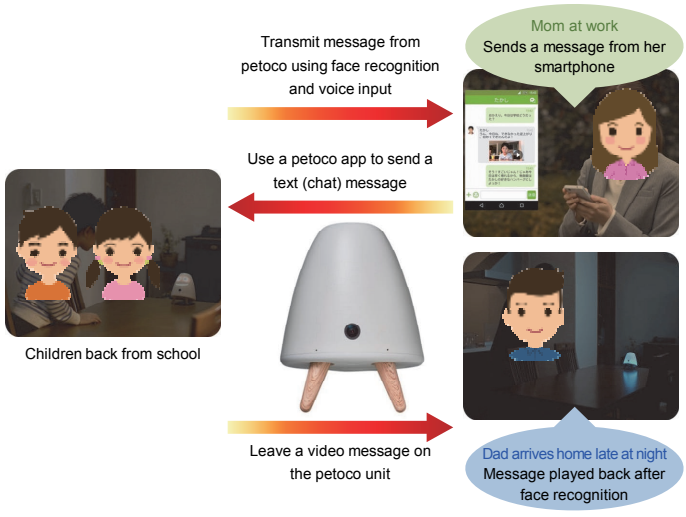


Figure 4 Basic function usage scenarios

Photo 2 Character display examples

Table 3 Main specifications

Size	Approx. 16×13×13 cm (H×W×D)
Weight	Approx. 355 g
Camera	Mounted on front
Speakers	Mounted on bottom
Microphone	Dual microphone
Image projection	Internal rear projection system
Lighting function	Full color LED
Wireless LAN	IEEE 802.11b/n/g
Bluetooth®※1	Ver. 4.0
OS Android™※2	OS 5.1
Power supply	DC 12V/2A (AC adapter included)
Technical features	Face recognition, speech recognition, echo cancellation, noise reduction

※1 Bluetooth®: A registered trademark of Bluetooth SIG, Inc. Used under license by NTT DOCOMO.

※2 Android™: A trademark or registered trademark of Google LLC.

projection system that projects images from inside petoco onto its smooth curved casing, which has a texture like frosted glass. To display uniform images on the product's curved surface, we used a focus-free[*6] Laser Beam Scanning (LBS)[*7] pico projector module[*8] (**Figure 5**). To enlarge the drawing area displayed on the housing, we increased the projection distance by using a mirror mounted inside the housing. A secondary effect of using a mirror is that it allows video to be projected without having to correct the left-right mirroring of the rear-projected images that would have occurred if they had been projected directly onto the interior of the housing (**Figure 6**).

3.2 Simple Operation

1) Voice Input and Output

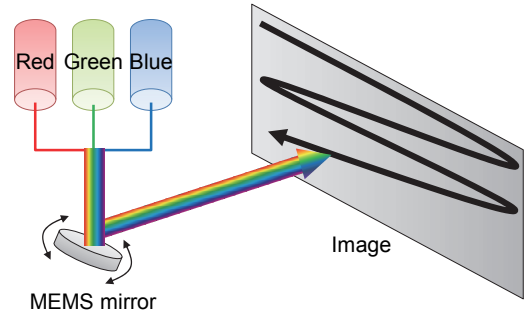
petoco is a device that is designed to be kept in the living room or dining room, and unlike handheld devices, it must always be capable of being operated from anywhere in the house. We must also assume it will be used by young children who may have difficulty inputting text. Based on these requirements, we decided to specialize petoco to operate only with voice input. This device uses a dual microphone to implement noise reduction[*9] and echo cancellation[*10] functions so that it can respond appropriately to voice input from users even in environments where there is background noise such as TV audio and other people talking. Furthermore, to provide feedback when voice input is received, the character responds in a way that corresponds to the function being performed so as to encourage feelings of attachment from children of lower elementary school grades at whom the product is targeted. During conversations with the device,

*6 Focus free: Focusing independently of the projection distance and angle, making focus adjustment unnecessary.

*7 LBS: A system where a semiconductor laser beam is scanned by reflecting it from a MEMS mirror in order to project images.

*8 Pico projector module: A module made with ultra-compact elements necessary for a projector such as an optical engine, driver and video processor.

*9 Noise reduction: Technology for suppressing noise. This is used in petoco to improve the accuracy of speech recognition in noisy environments.



MEMS (Micro-Electro-Mechanical Systems) mirror: A miniature machine incorporating a small mirror and drive unit

Figure 5 Overview of the laser scanning projector

A short projection distance results in a small display area. The image is also flipped horizontally.

A longer projection distance results in a larger display area. The mirror also corrects the horizontal inversion.

Figure 6 Using a mirror to perform back projection over a wider area

the character uses speech synthesis to generate different tones of voice.

2) Face Recognition

petoco must be able to recognize family members and provide each of them with the messages and other information they require. We implemented this capability by using a face recognition function. By registering the faces of family members the first time it is used, petoco can understand who is in front of it and can identify family members without any special operations.

Since this face recognition is always on, petoco can recognize family members at any time, say hello to them, congratulate them on their birthdays, and announce other pre-registered events such as anniversaries.

3.3 Communication Functions

1) Messaging

The messaging function received the most positive feedback from target users. It grew out of the existing idea of using sticky notes to pass messages between family members. A smartphone and petoco can exchange text, images and video, and petoco also has a function whereby people can leave messages by speaking directly to petoco itself. When

*10 Echo cancellation: A technique for removing echoes generated when a microphone picks up sound output by a device’s speaker.

messages are received by petoco, face recognition is used to identify who the message should be sent to, and by having the character read out the messages, petoco performs the role of a family hub in relaying communication. Also, when sending a message from a smartphone, this can be done easily by using a familiar text chat interface (**Figure 7**).

2) Video Chat

To implement video chats between petoco and smartphones, we used a technology called Web Real-Time Communication (WebRTC) that enables Peer-to-Peer (P2P)*11 communication between browsers to facilitate real time communication such as video calling. In addition to use within web browsers, we are also working on the provision of Software Development Kits (SDKs)*12 for various development environments to facilitate the development of software that works with petoco.

3) Monitoring

We also implemented a function that uses face recognition to send notifications to smartphones when family members arrive home. However, in the interviews, many people expressed concerns about the security considerations of this sort of feature. We therefore took care to ensure that there was no way of using the product to covertly monitor the home remotely via a smartphone, or for sending pictures from inside the home automatically when notifying a smartphone that someone has just arrived home.

3.4 Light Function

The enclosure contains RGB color-changing LEDs that allow it to light up. During the initial setting, it is possible to set a different color for each registered member of the family. When information has been stored in petoco, it can attract the recipient’s attention by blinking with the color corresponding to this individual (**Photo 3**). It can also be used as a table lamp.

3.5 Low-noise Design

It is assumed that petoco will be installed in living rooms and dining rooms where families spend most of their time. For this reason, one of our design policies was to avoid incorporating any sources of noise such as cooling fans into the

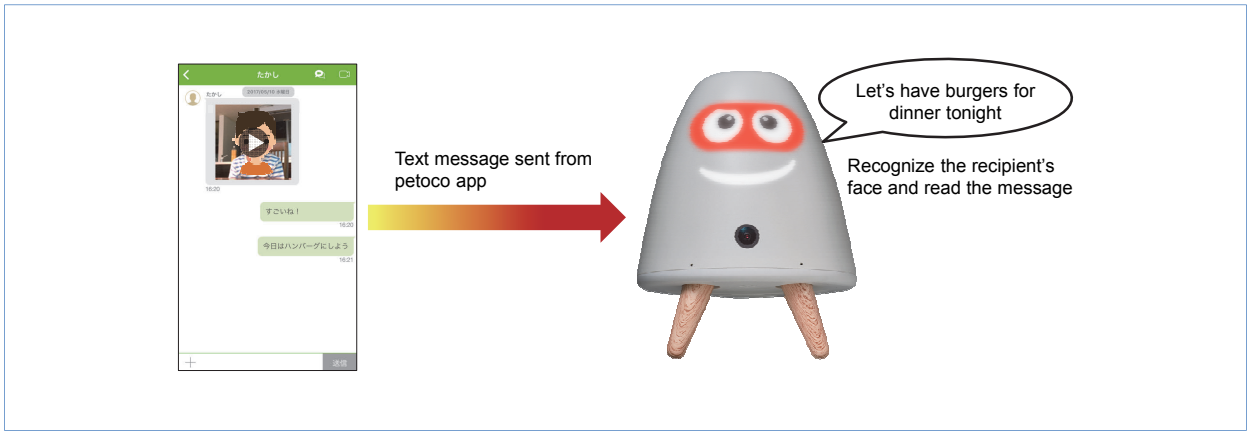


Figure 7 Transmitting and receiving messages

*11 P2P: A communication model where computers exchange information on an equal footing in contrast to the client-server model. In this article, exchanges of information between two mobile terminals and between a mobile terminal and petoco are all performed on an equal footing.

*12 SDK: A tool or set of tools used for software development.

Photo 3 Lighting example

product. Since this device uses a built-in projector to produce rear-projected images, our main concern was the dissipation of heat from the projector. Furthermore, we needed to ensure that the laser light from the projector could not escape and shine directly into people’s eyes, and that the light from the LEDs inside the enclosure could not leak out from the enclosure, which made it difficult to ensure a large enough opening in the enclosure to allow heat dissipation. We decided to study the use of heat dissipation technology using carbon fiber*13, which is not only more efficient than conventional technology based on aluminum or copper, but also performs better and is more reliable. After trying out various heat sinks, we confirmed that carbon fiber gave the best results, so we incorporated it into the device.

4. Conclusion

This article has described how we used design thinking to study concepts in the development of petoco, and how we embodied these concepts into a commercial product.

Although design thinking led us to an excellent concept, there are still many issues that must be resolved in terms of feasibility and cost in order to make this into a substantial device. We hope to continue this sort of new initiative in order to accumulate know-how and provide attractive new services.

*13 Carbon fiber: A material consisting of carbon atoms bonded together in long chains.

REFERENCES

[1] NTT DOCOMO: “Philosophy and Vision.” https://www.nttdocomo.co.jp/english/corporate/about/philosophy_vision/index.html

[2] Bureau of Statistics, Ministry of Internal Affairs and Communications: “Demographics, the Family Unit, and Other Changes in Social Structure,” 2018. http://www.soumu.go.jp/main_content/000452791.pdf

[3] N. G. Carr: “IT Doesn’t Matter,” Harvard Business Review, May 2003.

[4] H. Kikuchi, et al.: “The User-centered Design Process from Creation to Evaluation of a User Interface,” NTT DOCOMO Technical Journal, Vol.18, No.3, pp.30–37, Oct. 2010.

[5] T. Kelley and J. Littman: “The Art of Innovation: Lessons in Creativity from IDEO, America’s Leading Design Firm,” Doubleday Broadway Publishing, 2001.

[6] S. Oyama: “What is Design Thinking?” NIKKEI DESIGN, May 2014 (In Japanese).

[7] NTT DOCOMO Press Release: “Joint development of petoco: a home communication device,” May 2017.

[8] Cabinet Office Gender Equality Bureau: “Trends in number of double-employed households, etc.” 2014. http://www.gender.go.jp/about_danjo/whitepaper/h26/zentai/html/zuhyo/zuhyo01-02-08.html

[9] H. Ishii: “Fusion of Virtual and Real: Tangible Bits: User Interface Design towards Seamless Integration of Bits and Atoms,” IPSJ Journal, Vol.43, No.3, pp.222–229, Mar. 2002.

Technology Reports Intergenerational Communication Design Thinking Handwriting Devices

“Tomokaku”: A Handwritten Communication Concept

Communication Device Development Department Kei Ishiguro Nanako Mandai Takuto Tanaka
Mobile Division, DOCOMO Technology, Inc. Yuki Nishino
DOCOMO Innovations, Inc. Yusuke Saito

Since older generations and their children grew up in different eras, there is a gap between the ideas and tools that they use to communicate. This can make it difficult to achieve smooth remote communication between people of different generations. NTT DOCOMO is studying the use of design thinking as a means of resolving these issues. By focusing on handwritten communication, we have devised a concept called “Tomokaku®*1” and we have made a prototype that puts this concept into practice. In user evaluations performed using this prototype, the concept was highly rated by the users, confirming its market potential.

1. Introduction

In recent years, Japan’s population of elderly people (aged 65 and over) has been increasing, and the proportion of elderly people in the total population has also been increasing. On the other hand, as nuclear families become more common, fewer elderly people are living with their children, down from about 70% in 1980 to about 40% in 2015 [1]. Although the frequency of intergenerational communication within families is declining, NTT DOCOMO’s own surveys connected with this article have shown that the potential demand for this sort of communication is increasing, and that the growing tendency for elderly people to live separately is making it more important for them to communicate remotely by such means as telephones and email in addition to face-to-face conversations.

However, the means that people use to communicate with other family members who live

*1 Tomokaku®: A registered trademark of NTT DOCOMO, Inc.

elsewhere differs between generations. The majority of the elderly generation prefer to use voice communication via fixed and mobile telephones, and hardly ever use social networking services such as LINE*2 [2] [3]. Meanwhile, the same survey found that the children of the elderly generation mainly use LINE's chat function to communicate.

At NTT DOCOMO, we used the design thinking method to study different device concepts in order to resolve this intergenerational communication gap and facilitate smooth communication among families. As a result, by focusing on the User eXperience (UX)*3 of handwriting, which is an activity familiar to all generations, we proposed a new form of communication called "Tomokaku," which is modeled on the concept of a whiteboard. It has only a single function, but allows users to communicate both synchronously and asynchronously depending on the situation.

This article discusses the background of this concept, describes the details of the proposed "Tomokaku" UX, and presents the results of user evaluations performed using prototype devices.

2. Concept Study

A variety of different approaches can be considered for the study of methods for implementing smooth communication by eliminating the gap in communication means between generations within families, but here we performed a study based on the methods of design thinking [4] [5] that have recently been talked about as an effective method for creating innovation. Design thinking seeks to develop products that truly satisfy the needs of their users by focusing constantly on users rather than objects or techniques. This is achieved by repeating three processes: (a) gaining inspiration from users by observation, (b) creating ideas based on this inspiration, and (c) rapid prototyping and verification in order to evaluate these ideas.

In this study, we refined the concept by repeating the process of design thinking as shown in **Figure 1**. An overview of the representative points in each phase is presented below.

2.1 Inspiration: Extracting Insights from User Interviews

First, in order to observe and understand the users, we conducted in-depth interviews*4 with elderly users aged 70 to 80 years (the main target of this study) and users aged 40 to 50 years (the children of this generation). We also conducted similar interviews with extreme users*5 to gather hints on how to deal with issues from different perspectives.

These interviews included not only questions directly related to communication with distant family members, but also a wide range of other questions aimed at gaining a better understanding of the users, such as how they spend their time on an ordinary day. As a result of these interviews, we were able to gain insights*6 to help us further the aims of this study by resolving the intergenerational communication gap and facilitate smooth communication among families.

2.2 Ideation: Studying Concept Creation Principles Based on These Insights

For each of the insights gained in the interviews, we came up with multiple ideas for solutions to

*2 LINE®: A trademark or registered trademark of LINE Corp.
*3 UX: A general term for the experiences gained through the use or consumption of certain products or services.
*4 In-depth interview: An interview-based survey method where questioners explore the deep thoughts of the interviewees.
*5 Extreme users: Users with exceptional attributes and features who stand apart from ordinary target users.
*6 Insights: A user's essential desires and real intentions.

these issues, and we formulated concept creation principles by abstracting these ideas (Table 1).

2.3 Rapid Prototyping and Verification: Devising Concepts Based on Creation Principles

As an embodiment of our concept based on the

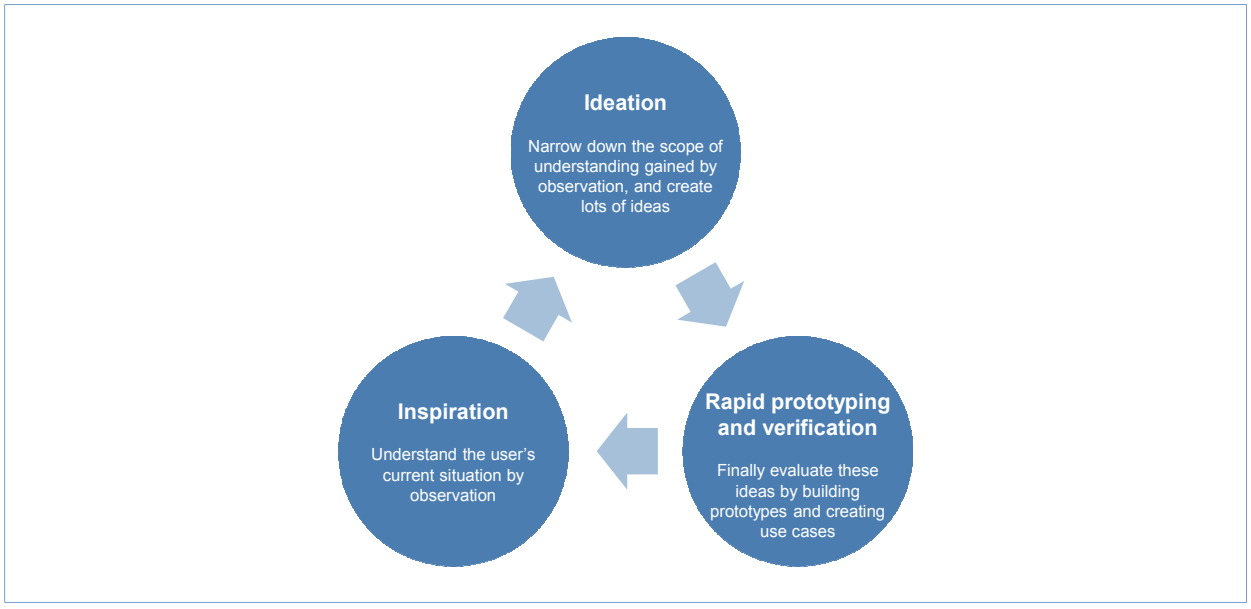


Figure 1 The design thinking process

Table 1 Our insights and the corresponding concept policies

Insights	Concept creation policy
When the elderly generation learns how to use new ICT equipment, they have difficulty in learning lots of new information at the same time, and this makes them feel uncomfortable with the deterioration of their learning ability.	Must be simple to operate without learning 1)
The elderly generation prefers physical things that they can manipulate by hand, because in digital systems they are unable to experience real objects.	Must present itself as an analog device that can be operated manually by physical interaction 2)
The elderly generation and their children want to communicate more with each other, but tend to concentrate on different points and different means of communication. For example, the elderly generation prefers synchronous modes of communication like the telephone, while their children prefer asynchronous modes of communication such as email.	Must have a means of choosing between synchronous and asynchronous communication according to circumstances 3)
The children of the elderly generation want to improve the quality of their parents' lives, and are positively disposed towards introducing ICT equipment for this purpose. However, they are hesitant to do so because they are concerned that this sort of technology is too complicated for the elderly generation, who will only be able to use it with a lot of support from their children.	Despite having simple functions, the device must be specified so that the children of the elderly generation can perform the required setup operations before presenting it to their elderly parents 4)

above principles, we concentrated on the UX concept of using handwritten communication as a familiar motif for users of all generations, and we implemented this concept by creating the rapid prototype shown in **Photo 1**. As a result, we finally arrived at the concept of Tomokaku, which involves remote handwritten communication via a whiteboard-type terminal.

Tomokaku is a very simple concept that allows users to share handwritten content on a whiteboard terminal in real time via the Internet. If one terminal is installed in the home of an elderly user and another in the home of this user’s child, this allows them to communicate just as if they are both writing on the same whiteboard (**Figure 2**).

The correspondence between the creation policy shown in Fig. 1 and the Tomokaku concept is as follows:

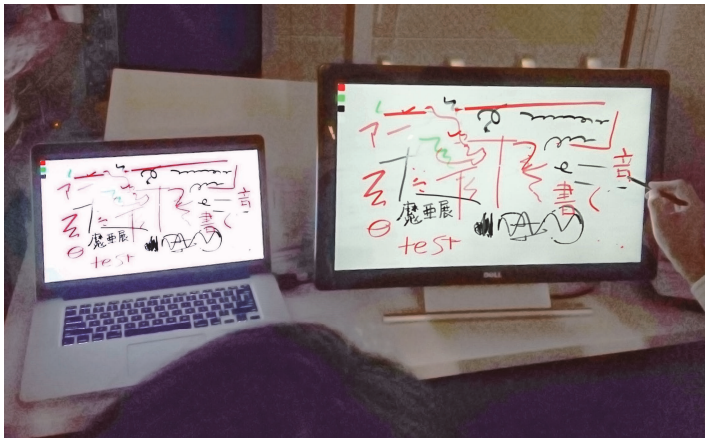


Photo 1　Rapid prototyping for verification

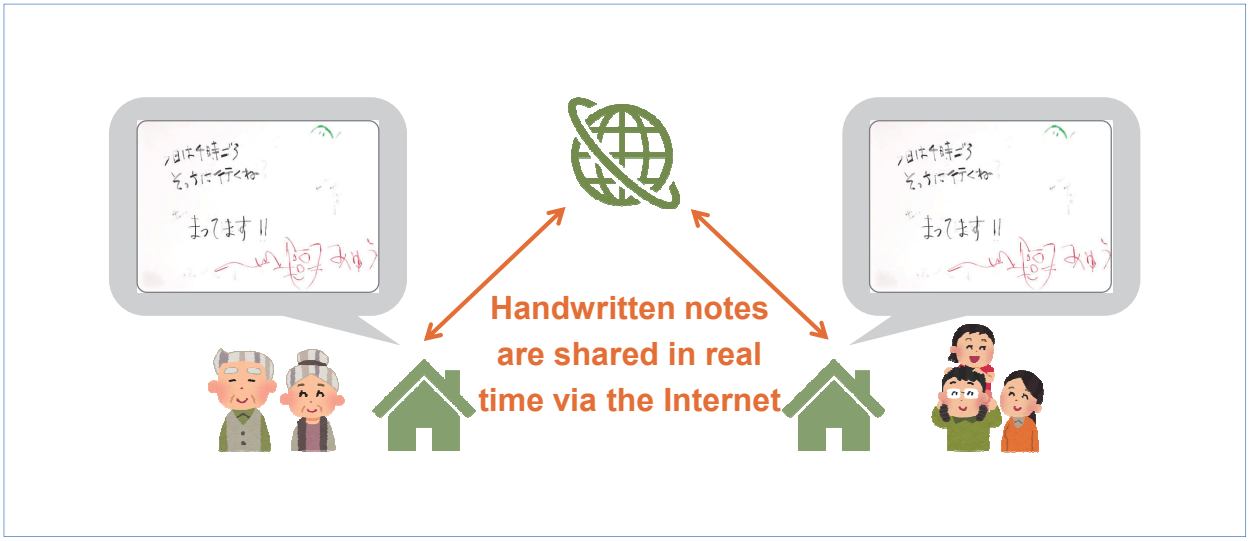


Figure 2　Overview of the Tomokaku concept

1) Must Be Simple to Operate without Learning

By concentrating on the simple function of using a pen, we made it possible to communicate by simple handwriting — an action that is familiar to all generations and does not require them to learn any new skills.

2) Must Present Itself as an Analog Device that Can Be Operated Manually by Physical Interaction

By making every effort to ensure that the product not only looks like a whiteboard but also functions like a whiteboard, we arrived at a design that gives users the real feeling that they are communicating through handwriting. Also, by having the external appearance resemble an ordinary whiteboard, we aim to avoid having it look like a “complicated piece of electronics” so as to eliminate psychological obstacles to using this product.

3) Must Have a Means of Choosing between Synchronous and Asynchronous Communication According to Circumstances

Since the handwritten input is just displayed synchronously on the receiving side, it is possible to choose between an asynchronous mode where the receiver can read the input when convenient like a written memo, and a synchronous mode where the pen input details and paths are displayed in real time on the receiving terminal (**Photo 2**), making it appear just as if the person at the transmitting side is writing directly on the terminal at the receiving side. Therefore, without users being particularly aware of it, a single function can be used to switch between synchronous and asynchronous communication as and when convenient.

4) Despite Having Simple Functions, the Device Must Be Specified so that the Children of the Elderly Generation Can Perform the Required Setup Operations before Presenting It to Their Elderly Parents

As described in 1) above, we made the overall functions simple. However, some functions such as initial settings can still seem complicated to elderly users. We therefore made it possible for children to perform the bare minimum of initial settings before presenting the terminal to their elderly parents. After completing the initial settings such as pairing the two terminals on the sending and receiving sides, the terminals can be used immediately, and the child can hand the terminal

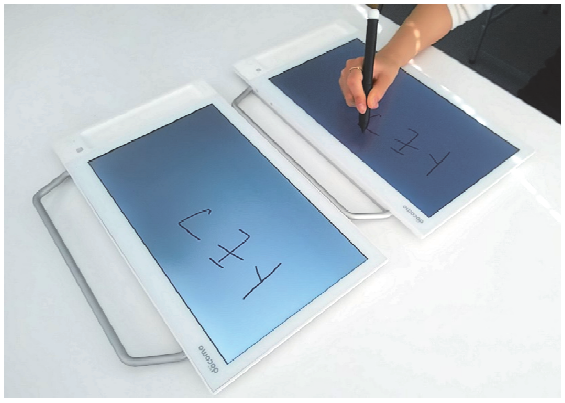


Photo 2 Handwriting entered via the sending terminal is displayed in real time on the receiving terminal

over without having to explain how to use it.

3. Prototyping

After several iterations of rapid prototyping, we performed technical studies and verifications to implement the abovementioned Tomokaku concept, and we made a prototype to evaluate its ultimate marketability. An overview of the prototype is presented below.

3.1 Setting Requirements Based on the Concept

It is important to set the requirements so that they embody the Tomokaku concept. The above concepts 2) and 3), which are related to hardware specifications, are particularly important for formulating specifications and must be studied carefully. The key points of the requirements we studied and formulated are as follows:

■ The terminal must take the form of an analog device with a physical feel that encourages manual operation (2.3 2))

(a) It must have the appearance of a whiteboard

- Its external appearance from the front must not include anything apart from the screen (i.e., no power buttons or the like)
- It must be capable of showing a full screen white display, without any icons or indicators
- The pen tip and shaft should resemble those of a whiteboard marker pen

(b) It must implement the same sort of functions and uses as an ordinary whiteboard

- As much as possible, the terminal must be as comfortable to use as an ordinary whiteboard (including high-precision detection of the pen tip, etc.) (1)
- It must not respond to the user's hand touching the screen (palm rejection) (2)
- It should not be necessary to switch the pen on, so it must always be in a state where writing can be performed immediately (3)
- When functions such as changing the pen color or using an eraser are provided, this must be done by using a different pen instead of navigating a menu system or the like (4)
- When using the pen to write on the terminal, it must emit the same sort of squeaking noise as a felt marker on a whiteboard
- The terminal must be battery powered so it can be easily carried around the home

■ Must have a means of choosing between synchronous and asynchronous communication according to circumstances (2.3 3))

- Although the terminals basically communicate asynchronously like handwritten notes, they must also have a function whereby the transmitting and receiving terminals update their displays synchronously in real time, in which case the receiving and transmitting terminals both emit squeaking whiteboard sounds when a user is writing on the transmitting terminal (allowing for synchronous communication if the user at the receiving end notices the sound)

We formulated the specifications of the prototype by considering the above requirements.

3.2 Selecting Technologies and Implementing Specifications to Realize These Requirements

This section describes the whiteboard-like pen operation and the implementation of the display screen, which were regarded as particularly important when formulating the specifications.

1) Simulating the Behavior of a Whiteboard Marker Pen

In embodying the requirements that we drew up for the prototype, the most important factor is to implement pen operations that resemble writing on a whiteboard. Table 2 presents a comparison of the main digital pen sensing technologies. As this table shows, the electromagnetic induction method*7 [6] is very effective for realizing this concept when considering benchmarks (1) through (4), which are listed as important requirements in the previous section.

We therefore adopted Wacom's EMR®*8 for this prototype, since it has the advantage of not requiring a battery and has already performed successfully in the drawing tablet market.

2) Implementing a Display Screen Like a Whiteboard

The abovementioned digital pen sensing technology and display technology are essential for detecting what is being written on the transmitting terminal and displaying it synchronously on the receiving terminal. For this prototype, we decided to use liquid-crystal displays, which are widely used in general-purpose products.

Furthermore, for the creation of a display screen with the appearance of a whiteboard, it is important to not only choose a display film that facilitates comfortable writing, but also to ensure that the background color is white. Therefore, it is assumed that the liquid crystal display is always completely white. However, in order to keep down the power consumption and prevent image burn-in, we incorporated a motion detection sensor and luminance sensor in the main unit so that the screen can be automatically turned off when there is nobody around or the surroundings are dark, thereby saving power when the equipment is not being used and at night.

Table 2 Comparison of digital pen sensing technologies

	Electromagnetic induction method	Capacitive sensing + Active pen	Capacitive sensing + Passive pen	Resistive sensing
(1) High detection accuracy	○	○	×	×
(2) Palm rejection	○	×	×	×
(3) No batteries required	△ (Only in some methods)	×	○	○
(4) Can change functions by using a different pen	○	△ (Only in some methods)	×	×

Capacitive sensing: Pen coordinates detected from a weak current change that occurs when touched by a human.

*7 Electromagnetic induction method: A method for the detection of coordinates by electromagnetic waves.
*8 EMR®: A registered trademark of Wacom Co., Ltd.

3.3 Outline of the Finished Prototype

Photo 3 shows the Tomokaku prototype. This prototype has a liquid crystal display and looks like a whiteboard with a handle. The main specifications of this prototype are listed below (Table 3).

4. Verification

4.1 Evaluation of Commercial Acceptability

To confirm the commercial acceptability of the proposed concept among its target users, we conducted user evaluations using the prototype described above. For this evaluation, considering the main target users of Tomokaku, we chose pairs of monitor users consisting of elderly people with poor IT skills (aged in their 60s and 70s) and their children (in their 40s) who lived in separate homes. They were asked to try out the terminals in their homes for about ten weeks (Photo 4).

For the duration of the verification test, the monitor users had prototypes installed in parts of their homes where they spend a lot of time, such as in their living rooms. Also, in order to evaluate the terminals in situations close to actual use, the users were given no particular instructions regarding when or how often to use them.

Photo 3 Our prototype Tomokaku device

Table 3 Key specifications of the Tomokaku prototype

OS	Windows 10 IoT Core
RAM capacity	1 GB
Storage capacity	8 GB
LCD size	14 inch
Resolution	1,366×768
Wireless LAN	Wi-Fi 802.11 a/b/g/n (2.4 GHz, 5 GHz)
Sensors	Luminance sensor, motion detection sensor
Battery capacity	1,500 mA
Input method	Electromagnetic induction (Wacom EMR)
Dimensions	385 × 267 × 16 mm

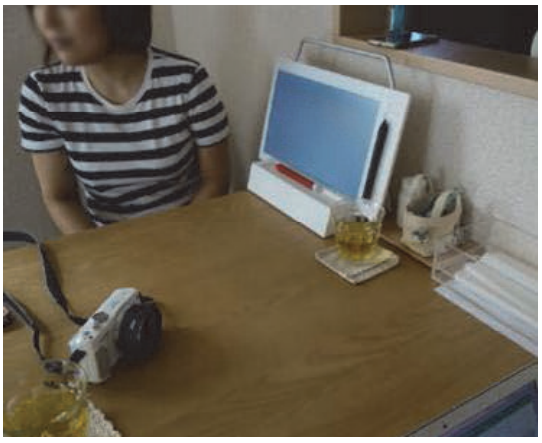


Photo 4 Evaluation at the home of a monitor user

4.2 Results and Discussion

After the trial period, we also interviewed the monitor users. They were asked to evaluate their overall satisfaction with the terminals with a score out of 10, and 80% of the users gave it high scores of 7 or more. To confirm the commercial acceptability of this concept, we also analyzed the comments from users in these interviews. The results are summarized for each concept creation policy below.

1) Must Be Simple to Operate without Learning

The users were highly appreciative of the fact that this prototype's single function and simple pen input method allow it to be used by family members of any generation, including grandchildren and the elderly people's own parents. On the other hand, some of the users with more experience in the use of ICT equipment were not as attracted to the simple single-function device of this concept, and did not see the need for dedicated equipment for this purpose. Therefore, it might also be effective to extend the target to include not only users who are unfamiliar with the use of ICT equipment but also users who are familiar with ICT equipment, allowing the tablet to run apps, for example, so that it can be used effectively according to the user's skill level.

2) Must Present Itself as an Analog Device that Can Be Operated Manually by Physical Interaction

Many users, especially those of older generations, were appreciative of the fact that Tomokaku provided them with a familiar means of communicating thoughts and ideas that might not seem important enough for a phone call or email.

3) Must Have a Means of Choosing between Synchronous and Asynchronous Communication According to Circumstances

The users liked the fact that Tomokaku shares handwritten notes directly, allowing the person at the receiving end to read finer nuances from the sender's handwriting compared with other asynchronous communication means.

Furthermore, Tomokaku was rated more highly as a synchronous communication means than as an asynchronous communication means, and the users particularly appreciated the feeling of closeness

they obtained through the function whereby the sender's handwriting is displayed in real time. Some users said it gave them a stronger feeling of interacting with a real live person, making them more inclined to respond quickly and express their true feelings.

4) The Device Must Be Specified so that the Children of the Elderly Generation Can Perform the Required Setup Operations before Presenting It to Their Elderly Parents

Since this prototype was designed to connect to the Internet via a wireless LAN, it was not possible for the children to perform the initial settings all by themselves, and their elderly parents had to perform some setting operations too. Some users seemed to have difficulty in completing these settings, so when studying refinements of this product for commercial use, we must consider specifications that can be set up more easily.

As stated above, we have confirmed that this concept — which was studied and proposed as a method that implements smooth communication by eliminating intergenerational gaps among families — is effective with regard to the insights, hypotheses and creation policies focused on at the study stage. Although there are still some issues to resolve, the proposed concept has been highly rated by users and appears to have strong market potential.

5. Conclusion

In this article, we discussed a handwritten communication concept called "Tomokaku" that was proposed for the purpose of resolving intergenerational communication gaps and facilitating smooth communication among families. On November 9–11, 2017, a prototype of this concept was exhibited at the docomo R&D Open House 2017 in Tokyo, where it was also well received and featured prominently in the visitor feedback questionnaires.

In the future, we will continue to study the commercialization of this design while addressing issues that came to light during the evaluation.

REFERENCES

[1] Cabinet Office: "2017 Aging Society White Paper: Aging and the implementation of countermeasures to old-aged society in FY2016," 2016 (In Japanese). http://www8.cao.go.jp/kourei/whitepaper/w-2017/zenbun/29pdf_index.html

[2] LINE website. https://line.me/en-US/

[3] NTT DOCOMO Mobile Society Research Institute: "Mobile Society White Paper: Trends in the use of smartphones and mobile phones to read data 2016–2017," 2016 (In Japanese).

[4] S. Oyama: "What is Design Thinking?," NIKKEI DESIGN, May Issue, pp.18–21, May 2014 (In Japanese).

[5] T. Kelley and J. Littman: "The Art Of Innovation — Innovation techniques learned from IDEO: the world's leading design farm," Hayakawa Shobo, 2002 (In Japanese).

[6] Wacom Corporation: "Electromagnetic induction type digitizer," (In Japanese). https://tablet.wacom.co.jp/what/news-img/W8002basis.pdf

Technology Reports

IoT eDRX Low-power UIM

Low-power IoT Technology —NTT DOCOMO Provides eDRX Technology and Low-power UIM—

Radio Access Network Development Department Masato Taniguchi

DOCOMO Technology, Inc. Packet Network Division Kazuki Takita Yuichiro Suzuki

Communication Device Development Department Takehiro Ida

In terms of battery life etc., there are demands for IoT communications terminals that consume less power than conventional terminals such as smartphones. In 2017, NTT DOCOMO began providing eDRX technology that reduces power consumption during standby and a low-power UIM to meet these demands.

This article describes the operating principles and characteristics of these technologies.

1. Introduction

Use cases of Internet of Things (IoT)*1 communications terminals could include devices such as environmental sensors or measuring instruments installed in locations where a power supply might not necessarily be secured. In such cases, power can be secured using batteries, although having many terminals installed in dispersed locations etc. will likely lead to cost blowouts due to the difficulty involved with replacing many batteries.

However, some IoT devices communicate less frequently than conventional devices, for example, they only send sensing data once a day [1]. In such communications scenarios, standby power consumption greatly affects terminal battery life.

Considering the above use cases and IoT device characteristics, NTT DOCOMO is developing low-power technologies for IoT to improve on battery life and power efficiency compared to regular LTE terminals such as smartphones. In 2017, DOCOMO developed and began supporting extended Discontinuous Reception (eDRX) low standby power communications technology, and providing a low-power

*1 IoT: A general term for a style of control and communication where various "things" are connected via the Internet or cloud services.

User Identity Module (UIM)*2.

Using sample devices, the results of measurements showed that the combination of the above two technologies can reduce power consumption by approximately 90% [2].

This article describes how these two low-power technologies are achieved on the DOCOMO network and in terminals.

2. eDRX Technology

To achieve lower terminal power consumption, NTT DOCOMO began supporting eDRX during standby (RRC_IDLE*3), as standardized in 3GPP Release 13 [3].

The following describes the eDRX operating principle, and the trade-off between power consumption and reception latency that should be considered in eDRX parameter settings.

2.1 The eDRX Power Saving Mechanism

Firstly, this section describes technology called Discontinuous Reception (DRX) that is used for energy-saving during standby with conventional terminals such as smartphones. When a terminal is on standby for incoming calls, it attempts to receive Paging*4 messages from the network at timed intervals (the DRX cycle). If a message addressed to the terminal is received at this time, the standby status is released, and communications with the network begin. **Figure 1** (a) shows a schematic diagram of the timing that the terminal attempts to receive Paging.

In this DRX operation, radio signal transmission and reception is suspended while no Paging reception is attempted, which reduces power consumption compared to attempting Paging reception continuously.

eDRX technology improves the effectiveness of

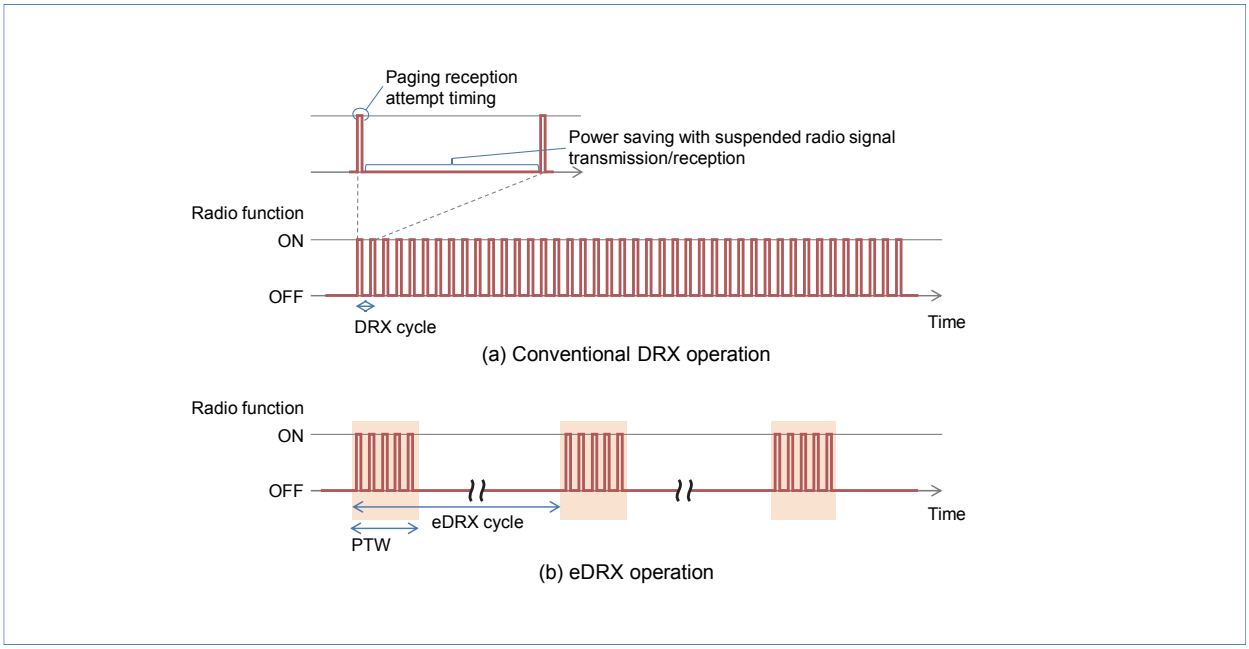


Figure 1 Comparison of DRX and eDRX paging reception timing

*2 UIM: An IC card storing subscriber information including the phone number and the IMSI (see *13), and inserted into the mobile terminal and used to identify the user.

*3 RRC_IDLE: A terminal status in LTE, in which the terminal context is retained in MME (see *5), but not retained in eNB. There are no data communications with RRC_IDLE.

*4 Paging: A procedure and signal for calling UE while camped in a cell in standby mode at the time of an incoming call.

power consumption reduction by reducing the occasions of Paging reception to even fewer than those of DRX described above. Fig. 1 (b) shows a schematic diagram of standby actions with eDRX.

Terminals operating eDRX only attempt to receive Paging according to the DRX cycle described above during periodic interval called the Paging Time Window (PTW). Setting parameters for the PTW length and cycle (eDRX cycle) determines the frequency of Paging reception actions, and hence determines the level of power saving.

For example, Fig. 1 (b) shows a DRX cycle of 1.28 sec, a PTW length of 6.4 sec (1.28 sec × 5), and an eDRX cycle of 102.4 sec (1.28 sec × 80). In this case, a terminal operating eDRX will attempt to receive Paging 5 times in the PTW during the 102.4 second period. Because regular DRX operations entail 80 receptions in the same period, by comparison, the number of reception operations with eDRX is greatly reduced.

2.2 Overview of eDRX Operations

1) Enabling eDRX and Parameter Settings

eDRX is a network-terminal coordination technology. For this reason, eDRX operating conditions (operation enabled/disabled, settings for eDRX parameters (eDRX cycle, PTW length)) must be established between the network and the terminal when the terminal starts etc. Network-terminal procedures are described below (**Figure 2**).

(1) From the terminal, an eDRX parameter request value is notified to Mobile Management Entity (MME)*5.

Waiting for eDRX to be enabled, the terminal requests operations enabling by setting eDRX parameters to an Attach*6/Tracking Area Update (TAU)*7 Request.

(2) The eDRX parameter values are determined by MME

When MME receives the Attach/TAU Request, it decides the actual setting values to be applied based on the eDRX parameter values notified from the terminal.

(3) eDRX parameter setting values are notified from MME to the terminal

MME notifies the terminal of the setting values determined in step (2) with Attach/TAU Accept, and the terminal applies the notified setting values to enable eDRX.

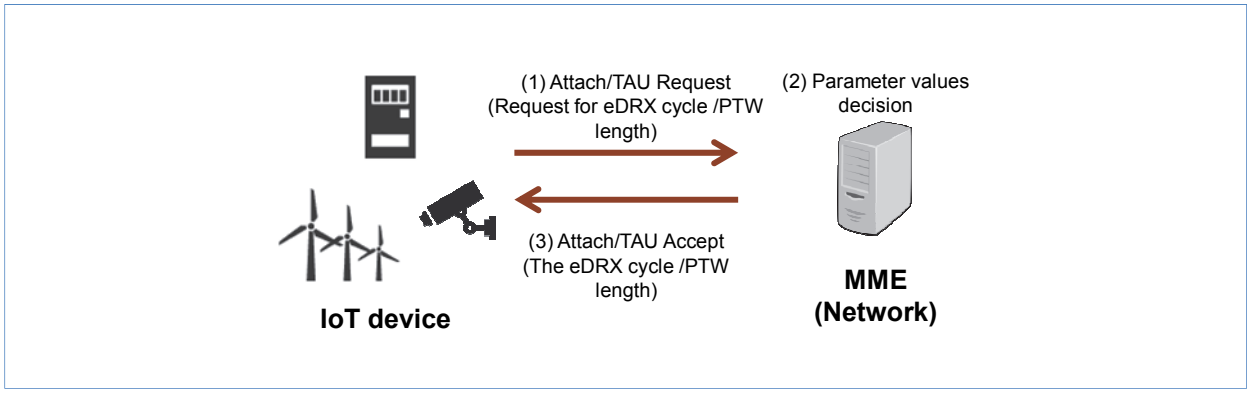


Figure 2 Enabling eDRX and setting parameters

*5 MME: A logical node accommodating a base station (eNB) and providing mobility management and other functions.

*6 Attach: The processing of registering a terminal with a network when terminal power is turned ON, or the state of being registered.

*7 TAU: Processing to update the terminal location registration information.

This procedure can be used to apply eDRX to only the terminals that require eDRX to be enabled, such as IoT devices.

2) Incoming Call Actions during eDRX Standby

Terminals on eDRX standby cannot receive Paging outside of the PTW interval (max. 43 minutes). However, users (e.g., IoT device administrators) or applications that need to communicate with the terminal send data to the terminal regardless of the terminal-side PTW timing.

To compensate for these timing differences, Evolved Packet Core (EPC)*8 uses High Latency Communication (HLCom)*9 technology to control transmission of Paging messages to the terminal with suitable timing, and buffer data for transmission.

Figure 3 shows the HLCom control sequence.

(1) When data packets from the server to the terminal arrive on the network, the arrival is notified to MME while the Serving Gate-Way (S-GW)*10 buffers the packet.

(2) If the terminal status is outside the PTW, MME derives the time until the terminal can receive the incoming packet in light of the time until the subsequent PTW opens and the Paging message response time etc. Then, MME notifies the packet buffering time (DL Buffering Duration) to S-GW, and S-GW extends packet buffering until the notified time.

(3) MME retains the Paging message transmission until the subsequent PTW open time, and then sends the Paging message to evolved NodeB (eNB)*11 at the PTW open time.

(4) After that, a communications bearer*12 is established between the terminal and the network, and the data packet is sent from S-GW to the terminal.

3) Dispersion of Paging Reception Timing between Terminals

While it's possible to call multiple terminals at the same time with one instance of Paging, the number of terminals that can be called at the same time is limited (a maximum of 16 in standard specifications). If calling to more than the maximum occurs at the same Paging instance, the network postpones Paging transmission. In cases where postponements are not possible due to extremely large amount of Paging conflicts, Paging messages are eventually discarded.

(a) Reception timing determination in DRX operations

To reduce conflicts of Paging with simultaneous timing, standard specifications prescribe dispersion of Paging reception timing for each terminal during the DRX cycle in conventional DRX operations. This dispersion is achieved by determining reception timing from a portion of digits from the terminal International Mobile Subscriber Identity (IMSI)*13, with a Modulo calculation*14 applied [4].

(b) Improving timing dispersion in eDRX operations

In addition to conventional dispersion methods, PTW open points are dispersed for each terminal in the eDRX function. The PTW open points are calculated by applying Cyclic Redundancy Check (CRC)*15 and Modulo calculations to SAE-Temporary Mobile Subscriber Identity (S-TMSI)*16.

In initial 3GPP discussions, extraction of PTW open points from IMSI was considered,

*8 EPC: The core network that accommodates radio access networks including LTE.

*9 HLCom: A function that controls Paging message transmission and buffers transmission data with appropriate timing for terminals that cannot respond immediately to connection requests using technologies such as eDRX.

*10 S-GW: A packet switch on the LTE Network for sending/receiving user data to/from P-GW.

*11 eNB: A base station for the LTE radio access system.

*12 Bearer: A logical user-data packet transmission path established among P-GW, S-GW, eNB, and UE.

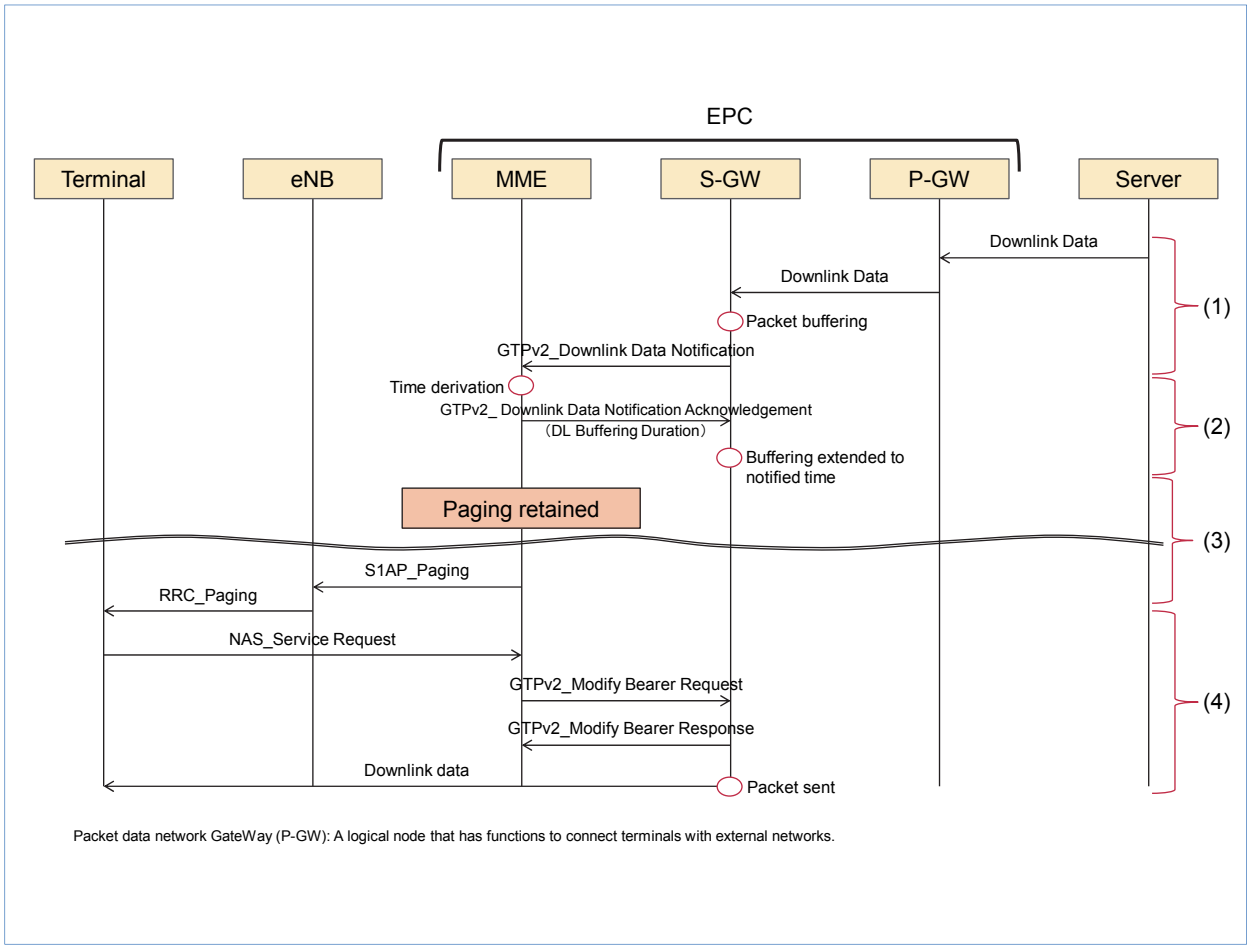


Figure 3 HLCom control sequence

however, to improve timing dispersion effectiveness, this was changed to determine PTW open points using S-TMSI instead of the IMSI used with conventional DRX timing.

Depending on the MME implementation, cases where the portions of the S-TMSI digits are close to being fixed can be considered. Thinking of this as a simple example of assignment in order of a series of numbers, the value becomes fixed across longer periods the higher the digits are. To ensure that PTW timing is dispersed appropriately in this case as well, CRC calculation is introduced, where a difference in a small part of the S-TMSI affects the entire part of the computation result used to determine PTW timing.

2.3 The Trade-off between Power Consumption and Incoming Call Latency in the eDRX Cycle

As mentioned, setting a longer eDRX cycle reduces the Paging reception opportunities, and thus

*13 IMSI: A number used in mobile communications that is unique to each user and stored on a UIM card.
*14 Modulo calculation: A calculation for obtaining the remainder when dividing one number with another.
*15 CRC calculation: A calculation which assumes the input bit string as a polynomial and obtains the remainder by division of the polynomial with a particular predetermined polynomial (the generating polynomial). Generally used for detecting errors that occur during data transmission.
*16 S-TMSI: Temporary numbers used for uniquely identifying users on the same network. Issued from MME.

reduces terminal power consumption. However, because it is not possible for a terminal to receive Paging outside the PTW, at maximum, an amount of time equivalent to the eDRX cycle could be required for a downlink incoming call to arrive at the terminal.

Considering the above characteristics, and depending on eDRX-compatible terminal users and system requirements, by setting the value of the eDRX cycle slightly lower than the maximum response latency allowed for incoming calls to the terminal, it's possible to achieve effective reductions in power consumption with eDRX while satisfying target latency requirements.

Moreover, when uplink data occurs to send from a terminal, the standby state is promptly released even if it is outside the PTW, and the uplink data transmission begins. Hence, there is no major impact on uplink data latency with the application of eDRX.

3. Low-power Consumption UIM

3.1 The Role of the UIM in Communications

The UIM operates with the power supplied from the terminal, contains subscriber information granted by the operator, and performs authentication procedures with the network through the terminal with location registration procedures. With these procedures, the subscriber information stored in the UIM is referenced to confirm that the terminal (UIM) is a legitimate user of the operator. In addition, an encryption key is created with the UIM in the authentication procedure to encrypt voice or data for communications with the network (**Figure 4**). Usage requires that the UIM is always

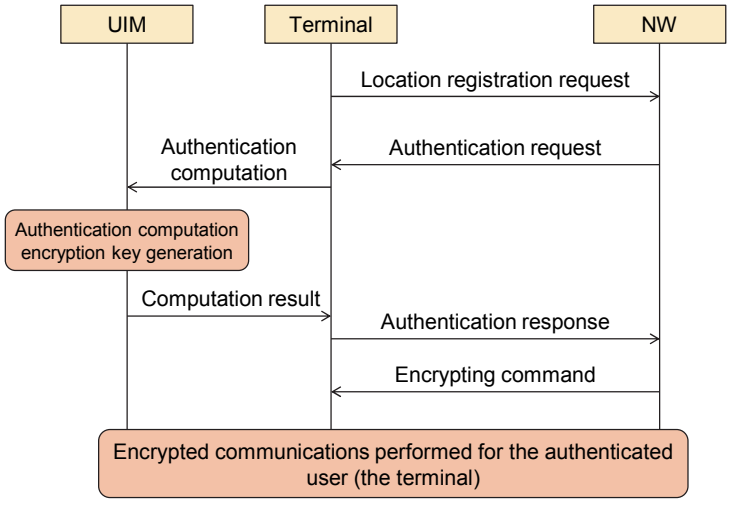


Figure 4 Location registration procedure

inserted into the terminal.

3.2 Power Saving Function

1) Suspension of Universal Integrated Circuit Card (UICC)*17 Presence Detection and Card Application Toolkit (CAT) Polling

As mentioned, because usage requires that the UIM is inserted into the terminal, there is a function called UICC Presence Detection to confirm insertion. With this function, the terminal periodically sends a STATUS command message to the UIM (at least once every 30 seconds in standard specifications), and confirms insertion with the received response. However, in Release 12 and later, this can be omitted to save power if there are no communications with the network. Nevertheless, when communications start, the terminal must send the STATUS command to confirm the presence of the UIM.

Also, to run applications loaded in the UIM, a procedure is periodically run via CAT polling signals from the terminal to confirm the presence of applications waiting to run in the UIM. If there are no applications to run in the UIM, it's possible to suspend polling the UIM in advance.

During the eDRX period, because the terminal is not communicating with the network, these two types of periodic communications between the terminal and the UIM can be suppressed.

2) UIM Deactivate Function

The UIM is constantly supplied power from the terminal so that it can respond to a request from the terminal at any time. However, if there are no communications (including location registration) with the network, the aforementioned UICC Presence Detection and CAT polling can be suspended so that communications between the terminal and the UIM do not occur. This method makes it possible to cut off the power supply to the UIM, reducing its power consumption to 0. Release 13 enables confirmation of whether suspension of the power supply to the UIM during communications is supported by UIM readout, so that the terminal can deactivate the UIM (suspend its operations) as required and cut off the power supply to it. However, because a UIM activate procedure and reading of several files in the UIM are required to start communications, which uses power, this technique is beneficial if a certain level of hibernation or above is expected.

According to the example in standardization documentation (3GPP TR31.970 Table2: Comparison of power consumption [5]) the current flowing to the UIM while the UIM and the terminal are not communicating is low at approximately 15 μA, although this adds up to around 0.36 mAh consumed per day. On the other hand, one activation (start of operations) of the UIM consumes approximately 0.06 mAh. For this reason, usage conditions with low frequencies of communications with the network (including location registrations) such as those of less than several times per day can be expected to be effective.

Combining these technologies enables suppression of power consumed by the UIM (**Figure 5**).

4. Conclusion

This article has described the operating principles and characteristics of the eDRX function and low-power UIM to achieve power saving with IoT terminals.

*17 UICC: An IC card used to record subscriber information such as telephone number and IMSI. UIM and SIM cards are used synonymously.

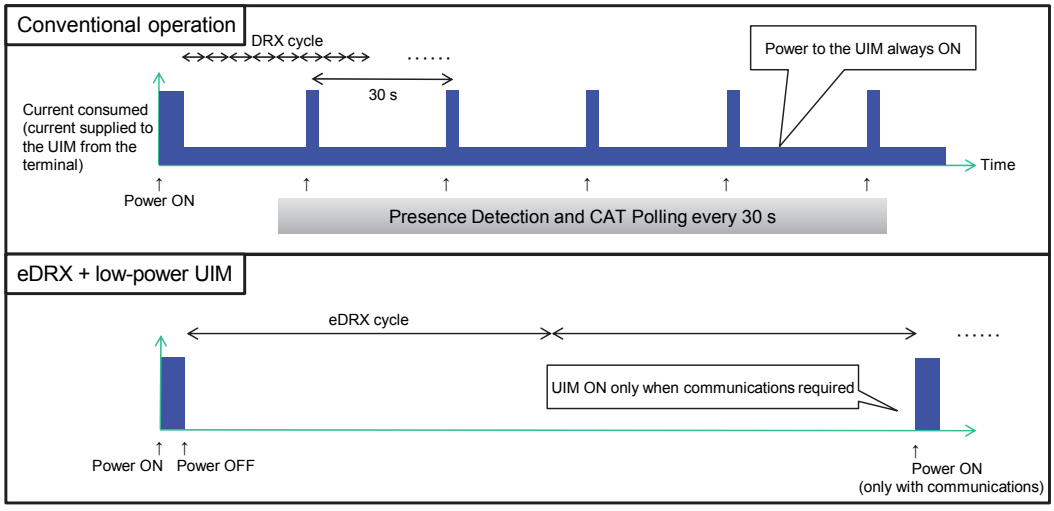


Figure 5 UIM operation and current consumption during standby

Going forward, NTT DOCOMO plans to realize optimized IoT communications environments through commercialization of communications technologies such as UE Category M1 (LTE-M) and UE Category NB1 (NB-IoT) [3].

REFERENCES

[1] 3GPP TR 23.887 V12.0.0: "Study on Machine-Type Communications (MTC) and other mobile data applications communications enhancements," Dec. 2013.

[2] NTT DOCOMO Press Release: "(Notice) NTT DOCOMO begins providing eDRX technology to reduce IoT communications device power consumption to approximately 1/5," Sep. 2017 (In Japanese). https://www.nttdocomo.co.jp/info/news_release/2017/09/26_01.html

[3] K. Takeda et al.: "New Technologies for Achieving IoT in LTE Release 13," NTT DOCOMO Technical Journal, Vol.18, No.2, pp.39-51, Oct. 2016.

[4] 3GPP TS 36.304 V13.8.0: "Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) procedures in idle mode," Dec. 2017.

[5] 3GPP TR31.970 V14.0.0: "Universal Mobile Telecommunications System (UMTS); LTE UICC power optimisation for Machine-Type Communication," Jun. 2017.

Collaboration Project

Investigative Research into Electromagnetic Interference with Medical Equipment from Radio Waves Emitted by Mobile Telephones and Smartphones

Research Laboratories Takahiro Iyama† Teruo Onishi

To promote the use of mobile telephones and smartphones in hospitals, NTT DOCOMO has been surveying electromagnetic interference on medical equipment due to radio waves emitted from these radio communications devices. Radio systems we surveyed are the 3G/LTE, 4G and WLAN systems supported by recent smartphones. We also surveyed candidate 5G frequencies.

This research was conducted jointly with Kanazawa University Hospital (Professor Keisuke Nagase, Department of Corporate Planning).

1. Introduction

During the spread of mobile telephones in the 1990s, there were heightened concerns about the potential for radio waves emitted by terminals electromagnetically interfering with other nearby electronic devices, in particular medical devices. Regarding these effects, the Electromagnetic Compatibility Conference Japan (EMCC)*1 surveyed first generation mobile communications systems (the analog NTT system), second generation mobile communications systems (Personal Digital Cellular (PDC)) mobile telephones, Personal Handy-phone Systems (PHS) and amateur radio etc. NTT DOCOMO was a key partner in proactively developing these research methods etc. [1]. In 1997, the Association published "Guidelines for the Use of Mobile Handsets to Prevent Electromagnetic Interference with Medical Electrical Equipment" in light of these research findings [2]. These guidelines prohibited bringing mobile telephone terminals into operating rooms or Intensive Care Units (ICUs) etc., and recommended switching mobile telephone terminal power off in laboratories, consultation rooms, patient rooms and treatment rooms etc., fundamentally prohibiting the use of mobile telephones etc. in hospitals. Furthermore, the Ministry of Internal Affairs and Communications conducted research on electromagnetic interference on medical equipment caused by third generation mobile communications

† Currently, Network Department

*1 EMCC: A joint industry-academic-government association concerned with issues such as electromagnetic interference.

system (3G) mobile telephone terminals including Wideband Code Division Multiple Access (W-CDMA), PHS and Wireless Local Area Network (WLAN) terminals, and published a report in 2002 [3]. It was determined from these findings that the continued application of the above guidelines was justified. Under those circumstances, individual hospitals formulated their own rules regarding the use of mobile telephones in hospitals upon making generalized judgments focusing on aspects such as mobile phone etiquette.

However, with the subsequent development of radio communications systems and smartphones etc., demands to use these devices in hospitals have increased from both workers and patients. This joint research was conducted in light of the latest conditions arising in medical equipment, and with the aim of understanding and analyzing the risks of electromagnetic interference with medical equipment. Specifically, in 2011, we conducted research on electromagnetic interference with medical equipment due to 3.9th generation mobile communications systems (Long-Term Evolution (LTE)) using 3G and its High Speed Packet Access (HSPA) standard and Frequency Division Duplex (FDD)*2 system, then from 2016 to 2017, conducted similar research on WLAN, fourth generation mobile communications systems (4G), in other words LTE-Advanced as LTE with Time Division Duplex (TDD)*3 system using the 3.5 GHz band, and candidate frequencies for fifth generation mobile communications systems (5G). This article describes an overview of this research and its findings.

This research was conducted jointly with Kanazawa University Hospital (Professor Keisuke Nagase, Department of Corporate Planning). To determine operational settings of various medical equipment and adverse effects due to electromagnetic interference, and to evaluate the degree of those effects, we received the cooperation of the clinical engineers in charge of operations, maintenance and inspection etc. of medical equipment in the hospital.

2. Overview of the Research

The following describes an overview of this research. Please refer to Reference [4] for details of external trends and various relevant standards.

2.1 Main Radio Waves Specifications

Table 1 shows the main specifications of radio systems and radio waves studied in this research. With 3G/FDD-LTE and TD-LTE, we used signals modulated in line with actual operating conditions for channels near the center of each frequency band assigned to NTT DOCOMO. For WLAN, we used signals modulated by selecting channels mainly

Table 1 Main specifications of radio systems and radio waves investigated

Radio system	3G/FDD-LTE	TD-LTE	WLAN (IEEE 802.11a, 11g)	5G (Candidate frequencies)
Frequency	800 MHz band, 1.5 GHz band, 1.7 GHz band, 2 GHz band	3.5 GHz band	2.4 GHz band, 5.2 GHz band, 5.6 GHz band	3.7 GHz band, 4.5 GHz band, 28 GHz band
Antenna	Half-wave dipole/terminal device			Half-wave dipole/horn

*2 FDD: A bidirectional communications mode that uses different frequencies and frequency bands for uplink and downlink communications.

*3 TDD: A bidirectional communications mode in which different time slots are allocated to uplink and downlink communications using the same frequency.

used with its frequency bands, and for 5G, we used unmodulated signals on 5G candidate frequencies. For the maximum transmission power, we referred to the technical standards for each radio system. Also, because it is known from previous research findings that electromagnetic interference with medical equipment occurs easily due to radio waves in intermittent mode repeatedly turning ON and OFF approximately 60 times per minute (see **Figure 1** (a)), we used signals to reproduce this. For the antennas to radiate the radio waves, we used a half-wave dipole antenna*4 and a horn antenna*5 that are more efficient than the built-in antennas in mobile telephones and smartphones and cause interference easily. In cases where interference occurred or if it was difficult to use these antennas, we used combinations of actual mobile telephones and smartphones for the corresponding radio system.

2.2 Research Site

Because of the necessity to avoid leaking of radio waves outside, we used a shielded room or a shielded tent set up in a room in the Kanazawa University Hospital.

2.3 Procedure

The most prone condition for electromagnetic interference with medical equipment are:

- Intermittent mode
- Max. transmission power
- Close contact with the antenna and medical equipment

Under the conditions above, we investigated impacts of electromagnetic interference on the medical equipment by scanning all over its all surfaces with a radio wave source (an antenna), and by changing the orientation of the antenna. This is called a basic investigation. When the medical equipment had sensors or cables attached, we also investigated interference effects on them. When electromagnetic interference occurred:

- We switched from intermittent modes to continuous mode (see Fig. 1 (b))
- We moved the antenna away from the medical equipment but maintained maximum transmission power
- We maintained close contact with the antenna and medical equipment, but lowered transmission power

to perform follow-up investigations. In cases where interference occurred, or it was difficult to use these antennas, we replaced the antenna with a terminal device. **Figure 2** and **Photo 1** provide an

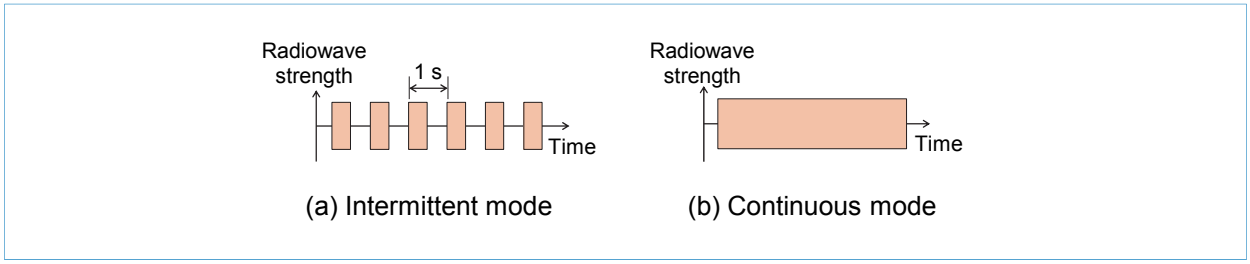


Figure 1 Transmission modes

*4 Half-wave dipole antenna: A very basic antenna which has two 1/4-wave length elements arranged in a straight line at the end of the cable (the feed point).

*5 Horn antenna: A type of antenna with a cone or pyramid shape, which emits a signal that is strong in a particular direction.

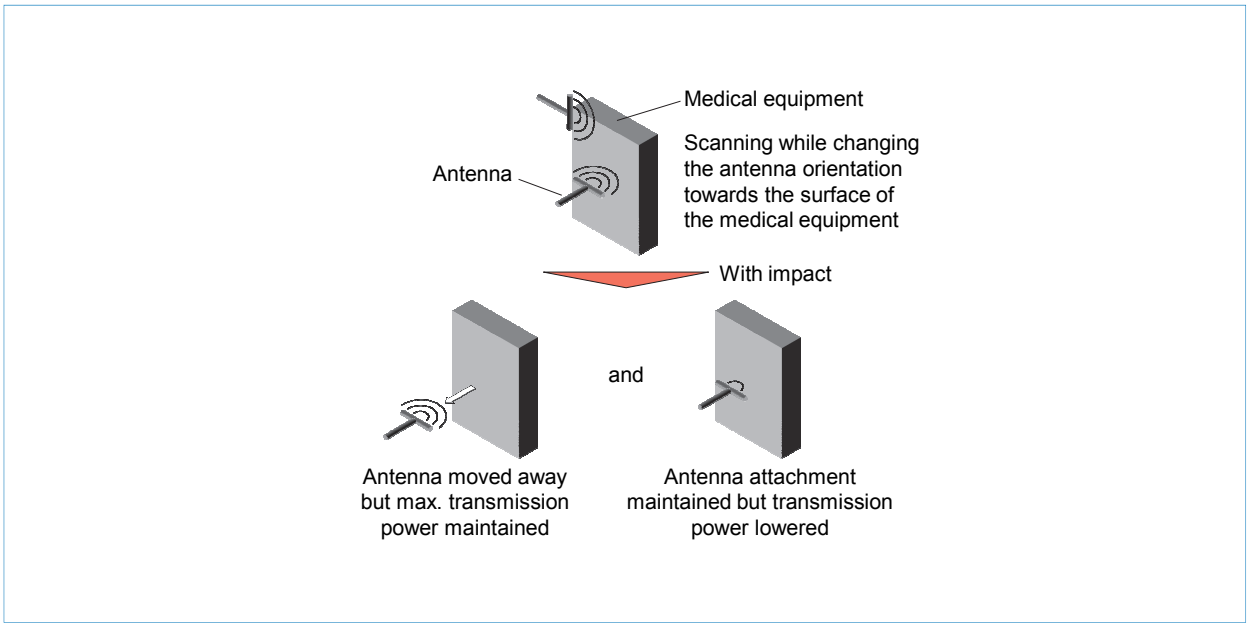
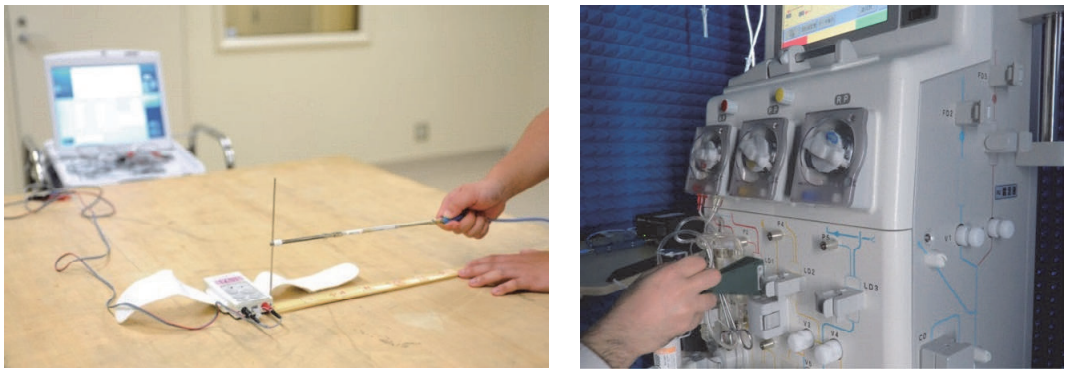


Figure 2 Investigation image



Investigating with a half-wave dipole antenna

Investigating with a horn antenna

Photo 1 The investigation scene

image of the investigation and the actual scene respectively.

2.4 Types of Electromagnetic Interference

When electromagnetic interference was confirmed, to determine the type of interference, we used the categories listed in a report published by the Ministry of Internal Affairs and Communications in 2002 as indices indicating the level of impact on actual medical treatment. These categories have been determined as physical types causing damage to medical equipment, and types adversely affecting medical treatment, as shown in Table 2. The higher the number of the category, the greater the impact on medical treatment.

Table 2 Categories of healthcare equipment failure

Physical types causing damage to medical equipment / Adverse impact on medical treatment	Normal	Reversible*1	Irreversible*2 (The following required for normal recovery)	
			Device operation	Device repair
No damage (normal)	1	–	–	–
Medical treatment disturbed*3	–	2	3	4
Misdiagnosis*4	–	3	4	5
Aggravated*5	–	4	5	6
Fatality*6	–	6	7	8
Catastrophe*7	–	8	9	10

Physical categories of medical equipment failure

*1 Reversible	The failure of the medical equipment can be reversed by moving the mobile phone that caused the interference away (or moving the medical equipment away).
*2 Irreversible	The failure of the medical equipment cannot be reversed by moving the mobile phone that caused the interference away (or moving the medical equipment away). The failure can only be rectified by human intervention or technical procedure.

Categories of medical treatment failure

*3 Medical treatment disturbed	Although the original medical purpose is maintained, the treatment does not go smoothly (small noise intrusions, baseline fluctuations, unpleasant sounds, character blurring etc.).
*4 Misdiagnosis	Medical equipment malfunction leads to wrong diagnoses and medical treatment errors. Although treatment is inappropriate, it does not put patient in danger (un-ignorable noise intrusions or baseline fluctuations, abnormal display values, stoppages due to alarms etc.).
*5 Aggravated	Medical equipment malfunction leads to medical treatment errors. If left unattended the malfunction could worsen the patient's medical condition (large changes in settings values, suspension of life support equipment, stoppages without alarms etc.).
*6 Fatality	Medical equipment malfunction leads to medical treatment errors. If left unattended the malfunction could cause patient fatality.
*7 Catastrophe	Medical equipment breakdown leads to major trouble. Patients could die or staff could be seriously injured.

2.5 Medical Equipment Investigated

The 2011 3G/FDD-LTE electromagnetic interference research included exhaustive investigations on a wide variety of medical equipment (53 devices in total) such as electric scalpels, infusion pumps, ventilators and ultrasonic echo devices in operating rooms, ICUs and laboratories in hospitals. In the following 2016 to 2017 investigation, and in light of the 2011 research results, electromagnetic interference was investigated using medical equipment thought to be susceptible to electromagnetic interference, which included 44 devices investigated with WLAN, 23 devices investigated with TD-LTE, and 23 devices investigated with 5G candidate

frequencies.

3. Findings

Table 3 shows a list of the maximum impact distance and maximum category for each radio system under the aforementioned easiest conditions to cause electromagnetic interference on the medical equipment. Similarly, **Figure 3** shows the maximum impact distance for each frequency band. The following describes findings for each of the radio systems and frequency bands.

3.1 3G/FDD-LTE

Testing exhaustively selected medical equipment showed that roughly 40% was affected by electromagnetic interference. The maximum impact distance was 80 cm, and the maximum category was 4. Medical equipment that is comparatively prone to electromagnetic interference included external cardiac pacemakers, electrocardiograms, nerve stimulators, infusion pumps and

Table 3 List of findings

Investigation	Radio system/ frequency band	No. of healthcare devices	Maximum impact distance	Maximum category
2011	3G/FDD-LTE	53	80 cm	4
2016	WLAN	44	28 cm	5
2017	TD-LTE	23	45 cm	4
	5G (Candidate frequency bands)	23	40 cm *No interference at 28.5 GHz	4

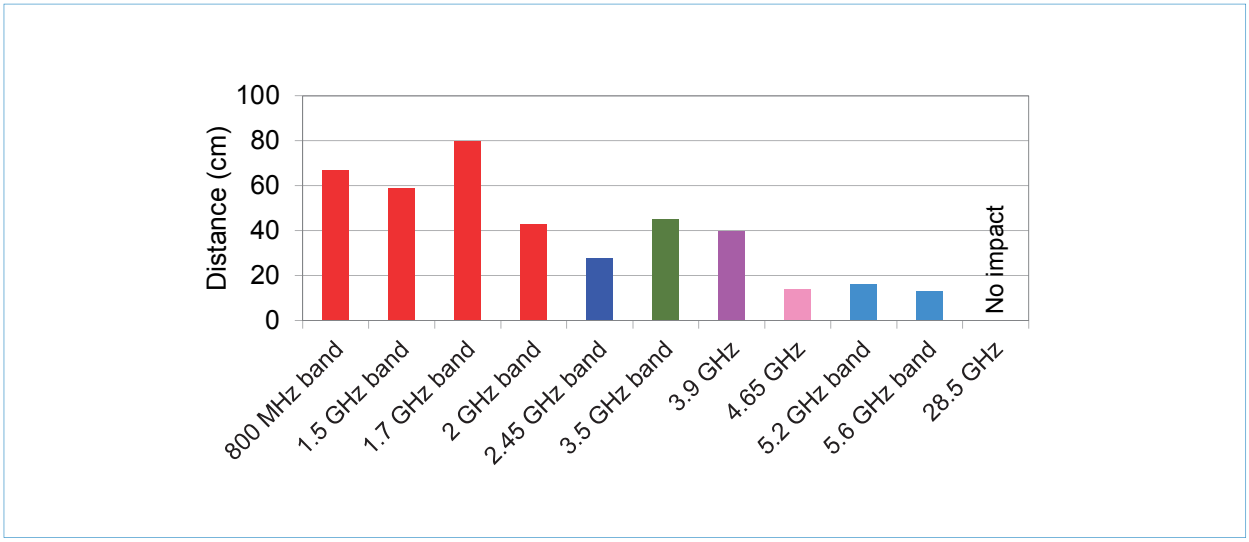


Figure 3 Maximum impact distance by frequency band

ventilators. We also confirmed that electromagnetic interference does not depend on radio wave Peak to Average Power Ratio (PAPR)*6 or modulation method, but is mainly dependent on radiated power [5]. This measured data was reflected in "Report on the Use of Mobile Phones and Other Devices in Hospitals" [6] issued by the EMCC in 2014. In light of past and most recent investigation findings, the Association has issued "Guidelines for the Use of Mobile Phones and Other Devices in Hospitals" [7] that include guidelines permitting use of mobile telephones etc. at distances of 1 m or greater from medical equipment according to appropriate rules created by hospitals.

3.2 WLAN (802.11a/g) and 3.5 GHz Band TD-LTE (LTE-Advanced)

Testing of medical equipment selected for its comparative proneness to electromagnetic interference showed that roughly 40% was affected by some form of electromagnetic interference. Resulting from WLAN and TD-LTE tests, the maximum impact distances were 28 cm and 45 cm, and the maximum categories were 5 and 4 respectively. However, with WLAN testing, even with actual terminals, only one medical device was affected by electromagnetic interference, which implies that the probability of electromagnetic interference from actual WLAN products is extremely small.

3.3 5G (Candidate Frequencies)

Testing of medical equipment selected for its comparative proneness to electromagnetic interference showed that roughly 20% was affected by some form of electromagnetic interference with the 3.7 GHz and 4.5 GHz bands. The maximum impact distances were 40 cm and 14 cm respectively, and the maximum category was 4. There was no electromagnetic interference caused in the 28 GHz band.

4. Conclusion

This article has described research into electromagnetic interference with medical equipment from radio waves emitted by mobile telephones and smartphones. The results of these investigations have contributed to the EMCC's guideline formulation and provided backing for their validity, and have formed the basis for relaxing the long-standing prohibitions on the use of mobile telephones in hospitals, hence enabling the use of mobile telephones and smartphones [6]. We have been investigating other radio systems including 5G, and in future we will contribute to achieving a world with ICT solutions to issues in the healthcare field by enabling safe and secure use of mobile telephones and smartphones in medical institutions.

REFERENCES

[1] T. Nojima et al.: "Electro-Magnetic Compatibility of Mobile Radio Communications," NTT DOCOMO Technical Journal, Vol.5, No.4, pp.6-14, Jan. 1998 (In Japanese).

[2] Electromagnetic Compatibility Conference Japan: "Guidelines for the Use of Mobile Handsets to Prevent Electromagnetic Interference with Medical Electrical Equipment," 1997 (In Japanese).

[3] Ministry of Internal Affairs and Communications: "Report on investigative research on the influence of radio waves on medical equipment etc." Mar. 2002 (In Japanese).

[4] S. Ishihara, J. Higashiyama, T. Onishi, Y. Tarusawa and K. Nagase: "Electromagnetic Interference with Medical Devices from Third Generation Mobile Phone Including LTE," EMC '14/Tokyo, 14P1-H6, May 2014.

*6 PAPR: Peak-signal-power to average-signal-power ratio. An index that indicates the peak signal power compared to the average signal power.

[5] K. Nagase, S. Ishihara, J. Higashiyama, T. Onishi and Y. Tarusawa: "Electromagnetic interference with medical devices from mobile phones using high-speed radio access technologies," IEICE ComEX, Vol.1, No.6, pp.222-227, 2012.

[6] Electromagnetic Compatibility Conference Japan: "Report on the Use of Mobile Phones and Other Devices in Hospitals," Aug. 2014 (In Japanese).

[7] Electromagnetic Compatibility Conference Japan: "Guidelines for Use of Mobile Phones and Other Devices in Hospitals - For Secure, Safe Use of Wireless Communication Devices in Hospitals -," Aug. 2014.

Topics

Low-power LTE Terminal for Gas Smart Meters

Solution Service Department Tsuyoshi Nagasaka Masanori Ishida
Shogo Yabuki Masashi Onogawa

Gas smart meters are fundamental platforms that improve consumer services by enabling visualization of energy usage, emergency valve shutoff, alarm notifications and flexible pricing. Also, since they solve issues with meter reading in situations such as those where meter reading workers cannot enter apartment buildings due to automatic locking, and meet gas supplier needs by bringing efficiency to their work and rationalizing facilities construction to reduce costs, implementation of gas smart meters is ongoing [1].

Currently, new technologies called "U-bus" and "U-bus air" standardized by the Japan Utility Telemetering Association (JUTA) are being introduced in the town gas industry [2].

- Enabling packet communications with higher speed than the conventional communication specifications, U-bus is a next-generation interface that connects gas meters, sensors and gas devices such as alarms and burning appliances easily to a center system*1. This technology promises improved services (energy usage visualization, remote valve opening and closing).
- U-bus air is short distance radio communications operating in the 920 MHz band that enables relaying between gas meters using multihop communications*2. The system makes it possible to connect gas meters and so forth that are outside the coverage area of a Wide Area Network (WAN).

To achieve the gas smart meter system, a WAN device to connect to a network configured with the aforementioned U-bus and U-bus air technologies, and a WAN to connect to the center system are required (**Figure 1**).

WAN devices are anticipated for their 10-year operating durability the same as other gas devices, as well as their ability to relay data from multiple devices through U-bus and U-bus air. However, because of the difficulties in providing stable power sources with outdoor installation etc., these devices must run on batteries, which means if they are to operate for at least 10 years, the communications systems used with the WANs must be energy-saving.

Also, to provide stable and secure services to general consumers, not only requirements for uplink communication to send the meter data to a

*1 Center system: Equipment to manage devices such as gas meters and acquired data.

*2 Multihop communication: A communication system that enables terminals to communicate directly with each other, or enables widely separated terminals to exchange data by relaying via multiple terminals on a network of hierarchically arranged terminals.

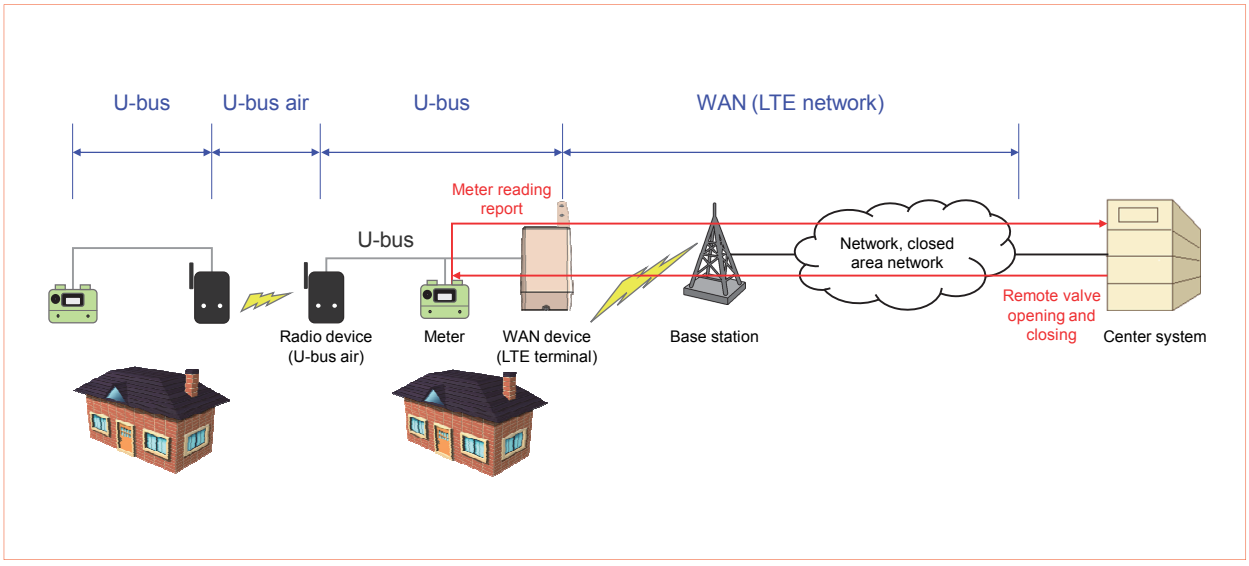


Figure 1 Gas smart meter system overview

center system, but also requirements for downlink communication to remotely open and close meter valves etc., must be satisfied. Therefore, low-power, bidirectional communications with high communications success rates must be achieved.

1) Solving Issues with LPWA Technology

Technology called Low Power Wide Area (LPWA) is gaining attention as a low-cost solution for the aforementioned issues. In general, Internet of Things (IoT)*3 devices differ from communications terminals such as smartphones, because the frequency of uplink communications from terminals is low (around 1 time per day), and the amount of data sent in a single transmission is extremely small. LPWA is IoT device data communications technology that covers a wide area with low power consumption instead of small amounts of transmitted and received data, and satisfies the requirements for gas smart meters.

With the demand expansion in the IoT field including gas smart meter projects, the 3rd Generation Partnership Project (3GPP) standardization organization newly prescribed specifications for IoT devices in LTE Release 13 formulated in 2016. One of the LPWA technologies, extended Discontinuous Reception (eDRX), is a new technology prescribed in Release 13 [3]. Conventional DRX is technology that uses intermittent signal reception, and saves power by putting terminal communications functions to sleep during periods of no reception. eDRX promises even greater power savings by significantly extending the intermittent reception interval (**Figure 2**).

2) Overview of Low-power Communications Terminals for Gas Smart Meters

Based on the stable communications platform that NTT DOCOMO has provided to date, we have developed an LTE terminal (a WAN device) that supports eDRX low-power technologies for gas smart meters. **Photo 1** shows the external appearance of the terminal.

The LTE module in this LTE terminal is designed with the assumption that the LTE terminal will still only communicate small amounts of data, and adopts Category 1, which promises lower power consumption and chip cost reductions compared to UE Category*4 used with smartphones. Using an 81.92 sec intermittent reception interval (eDRX

*3 IoT: Mechanisms of mutual control via exchange of information among various "things" connected to the Internet.

*4 UE Category: An index indicating the communications performance of a terminal.

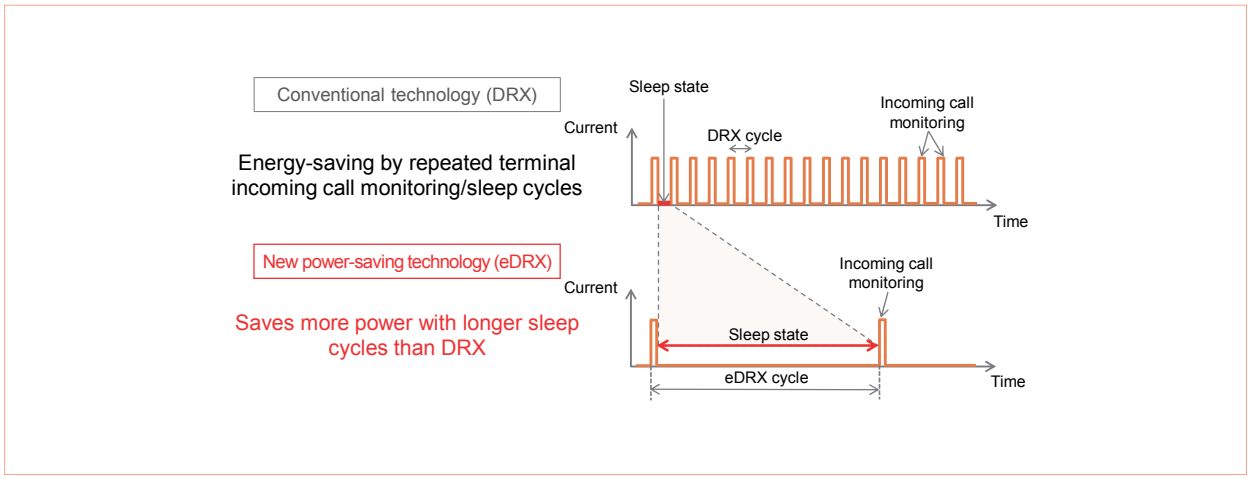


Figure 2 eDRX overview

Photo 1 LTE terminal external appearance

applied) in a prototype using the newly developed low powered User Identity Module (UIM)*5, we were able to confirm more than 90% power saving compared to a conventional UIM with an intermittent reception interval of 1.28 sec (eDRX unsupported) [4]. This system promises operations of 10 years or more with batteries, and cheaper costs through reduction in the number of batteries required.

After installation, one LTE terminal issue is function updating, because it has been necessary to collect the terminals for updating, which affects operations.

To solve this issue, we equipped the terminals with a function to update firmware from the center system (remote firmware update function). This enables flexible function addition and alteration to expand the functionality of gas smart meter systems.

Also, a smooth batch terminal control from the center system is enabled by acquisition of terminal in-range information etc. and updates are dispersed geographically due to concerns about increased traffic and increased communications time, and the associated increased power consumption in areas with a high density of terminal installations (**Figure 3**).

3) Future Approaches

Working towards automating the meter reading work of town gas suppliers, we are evaluating the performance and validity of these eDRX communications terminals and communication systems from FY 2017 to FY 2018. This specifically entailed installing eDRX-enabled terminals and test gas meters in the houses and apartments of 50 households in FY 2017, and 500 households in FY 2018. Using an actual LTE network, we evaluate communications success rates, time required for communications and power consumption [5].

The UE Category called Category M1, Category NB1 is newly supported in 3GPP Release 13 [3]. These technologies promise to reduce LTE module

*5 UIM: A card on which information for identifying a subscriber is recorded.

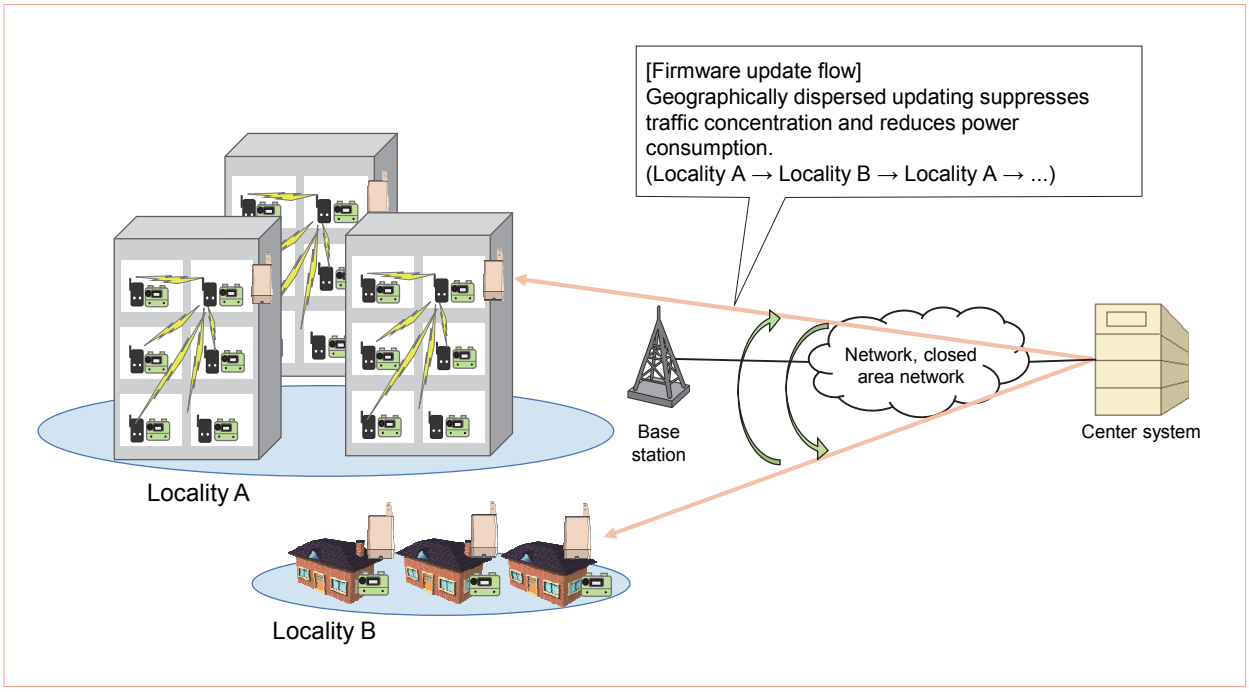


Figure 3 Firmware update overview

power consumption and costs by restricting the terminal transmission/reception bandwidth. In the future, we plan to study and verify Category M1 and Category NB1 in parallel with Category 1 field trials to further reduce power with terminals using these technologies, and lower costs by reducing the number of required batteries.

Also, in industries other than the town gas industry, customer needs for communications terminals for smart meters vary greatly. For example, in the LP gas industry, it is necessary to replace gas cylinders before the gas remaining in the cylinder runs out, and the replacement cycle relies on the experience of meter reading workers. However, using LPWA technology, information about the amount of gas remaining in gas cylinders can be collected in real time. Then with AI and this collected information, replacement cycles can be optimized and thus reduce workloads through greater efficiency. In this way, we are not only aiming for stable operation of gas smart meters currently on offer, but also taking initiatives to solve a range of issues facing customers.

REFERENCES

[1] Ministry of Economy, Trade and Industry, Agency for Natural Resources and Energy, Electricity and Gas Industry Department: "The current state and issues surrounding smart meters," May 2010 (In Japanese). http://www.meti.go.jp/committee/materials2/downloadfiles/g100526a04j.pdf

[2] NPO Corporation Japan Utility Telemetering Association: "New telemetering communications methods - U-bus, U-bus air standardized -" (In Japanese). http://www.teleme-r.or.jp/u-bus/outline/contents.html

[3] K. Takeda et al.: "New Technologies for Achieving IoT in LTE Release 13," NTT DOCOMO Technical Journal, Vol.18, No.2, pp.39-51, Oct. 2016.

[4] NTT DOCOMO Press Release: "(Notice) NTT DOCOMO begins providing eDRX technology to reduce IoT communications device power consumption to approx. 1/5," Sep. 2017 (In Japanese). https://www.nttdocomo.co.jp/info/news_release/2017/09/26_01.html

[5] NTT DOCOMO Press Release: "(Notice) NTT DOCOMO develops low power eDRX technology-compatible LTE terminals for gas smart meters, and begins field trials," Nov. 2017 (In Japanese). https://www.nttdocomo.co.jp/info/news_release/2017/11/09_00.html

Medal with Purple Ribbon in 2018 Spring Medals of Honour

In 2018 Spring Medals of Honour, Seizo Onoe, a former NTT DOCOMO Executive Vice President, Member of the Board of Directors and Chief Technology Officer (CTO)† received the Medal with Purple Ribbon for his development of radio access control technologies for 3rd and 4th generation mobile telephone systems.

The Medal with Purple Ribbon is awarded to individuals for their inventions and discoveries in the science and technology fields, or their outstanding performance in the academic, sports, arts and culture fields.

Although implementation of mobile communications began in 1979 in Japan, with the expanded usage of computers and the Internet that followed in the 2000's, demands arose for high-speed data transmission and improved frequency usage efficiency with 3rd generation mobile communications systems (hereinafter referred to as "3G") to enable mutually compatible usage through the same systems around the world. After that, in the 2010s, demands arose to achieve 4th generation mobile communications systems (hereinafter referred to as "4G") and smooth transition to 4G for even greater speed and capacity as mobile telephone systems became more entrenched as social infrastructure and traffic continued to increase.

To meet these demands, Mr. Onoe developed new radio access control technology for the 3G international standard IMT-2000 (W-CDMA) to enable flexible base station area configuration. Then, he developed a mechanism for the 4G international standard IMT-Advanced (LTE/LTE-Advanced) to ensure continuity of communications with the shift between the different generation base stations (between 3G and 4G).

These developments contributed to the achievement and spread of 3G and 4G both in Japan and around the world and have enabled comfortable communications with smartphones etc., thus contributing to improving convenience in the lifestyles of citizens. Mr. Onoe received this award in recognition of these achievements.

Mr. Onoe is in the center of the photo

† Currently, President and CEO of DOCOMO Technology, Inc.

NTT DOCOMO
Technical Journal Vol.20 No.2

Editorship and Publication
NTT DOCOMO Technical Journal is a quarterly journal edited by NTT DOCOMO, INC. and published by The Telecommunications Association.

Editorial Correspondence
NTT DOCOMO Technical Journal Editorial Office
R&D Strategy Department
NTT DOCOMO, INC.
Sanno Park Tower
2-11-1, Nagata-cho, Chiyoda-ku, Tokyo 100-6150, Japan
e-mail: dtj@nttdocomo.com