NTT DOCOMO

Technical Journal

vol.18 No.3 | Jan.2017

DOCOMO Today

Technology Reports

NTT docomo

DOCOMO Today

Bolstering the Intellectual Property Power of NTT DOCOMO

Takashi Komoro†

General Manager of Intellectual Property Department

How well known is the intellectual property power of NTT DOCOMO? The company has been deeply committed to Research and Development (R&D) over many years, and one result of this work has been the steady acquisition of patents. NTT DOCOMO currently holds about 14,000 patents worldwide. However, if we were to ask if these assets were being sufficiently used for the benefit of business, we would unfortunately have to say that more effort is needed.

Mobile communications technology represents a major strength of R&D at NTT DOCOMO—it is the company's main field that accounts for more than 40% of its patents. As you know, international standards such as W-CDMA and LTE have been established in mobile communications, and such patents that are deemed necessary to manufacture products complying with these standards are called "standard essential patents." Patents, in general, are frequently locked within a company's proprietary technology and used as such, but a company is obligated to license any standard essential patents it might hold to anyone under fair and reasonable terms to promulgate the use of standards. This mechanism, which has been in place since 1994, has enabled a variety of firms to manufacture telecommunications equipment leading to overall market expansion. Since 2009, however, lawsuits related to standard essential patents as in the case of "Apple versus Samsung" have surged, and the need for measures that can achieve some sort of balance here has been felt to enable further expansion of the telecommunications market.

Against this background, NTT DOCOMO has been making major contributions to the development and standardization of new telecommunications technologies, and it has come to hold many standard essential patents for 3G, 4G, and other standards as a result. In a field dominated by global telecommunications manufacturers, NTT DOCOMO has consistently held a top share in such patents [1], which is quite unique as an operator and a Japanese firm. Our Intellectual Property Department aims to license and monetize NTT DOCOMO's standard essential patents throughout the world, and to this end, it has been participating in patent pools and negotiating licensing terms with individual firms. Moreover, in recent years, it has been putting a great deal of effort into negotiations with global handset manufacturers, which has begun to show results, if only a few so far [2]. At present, the submission of 5G-related patent applications is well underway, and to increase NTT DOCOMO's share of standard essential patents for 5G and expand licensing activities, we are ramping up our support for inventors.

At the same time, service-related technologies as in NTT DOCOMO's Smart Life initiative are becoming increasingly important. In contrast to standard essential patents, the company is using these technologies as a weapon for achieving business flexibility by preventing imitation by other companies and using them as a basis for alliance forming. For this reason, an effective approach is to establish rights to strong points of NTT DOCOMO services as strategic patent clusters, even if they are not necessarily innovative ideas that could lead to the creation of new business areas. In "mobile spatial statistics," for example, the Intellectual Property Department has teamed up with departments of concern in the initial stages of service planning so that strategic patenting can be promoted in fields heavily affected by imitation by other firms in various types of applications. From here on, there will be increasing collaboration with diverse partners as part of NTT DOCOMO's "+d" initiative, but to make good use of ideas and technologies unique to NTT DOCOMO while fully protecting them, I would like to see the Intellectual Property Department form even closer ties to concerned departments at early stages of service planning and to work together as one from the creation of intellectual property to its use.

Returning to intellectual property power, NTT DOCOMO is creating an extensive amount of intellectual property with great potential in the above ways. However, the possession of intellectual property by itself cannot contribute to business since they incur maintenance costs such as patent annuities. In the Intellectual Property Department, we are taking up a variety of challenges to intensify the use of patents to enhance NTT DOCOMO's intellectual property power in a true sense. Nevertheless, whatever actions we take will not function well if we do not have a close relationship with the inventors and departments of concern. We will therefore make every effort to use intellectual property in a way that contributes to business while working with all concerned in a positive and supportive manner.

References

[1] Cyber Creative Institute: "Evaluation of patents declared to ETSI as essential to LTE related standards."
http://www.cybersoken.com/en/research/lte/

[2] NTT DOCOMO Press Release: "DOCOMO Grants Patent License to HUAWEI," Jun. 2016.
https://www.nttdocomo.co.jp/english/info/media_center/pr/2016/0610_00.html

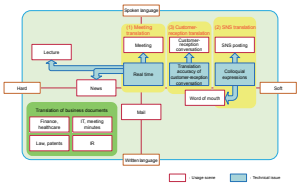
† Currently NTT Software Corporation, Senior Vice President, General Manager, Strategic Business Integration Division

NTT DOCOMO
Technical Journal
Vol.18 No.3 Jan. 2017

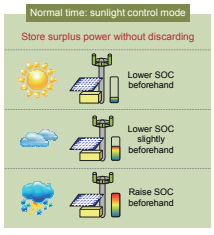
Contents

DOCOMO Today

Technology Reports



(P.12)



(P.23)

News

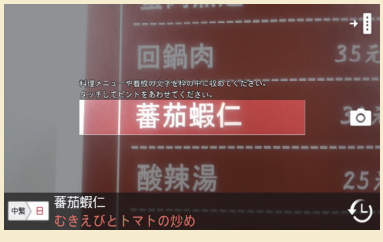
(P.41)

Hanashite Hon'yaku

Jspeak

Mail Hon'yaku

Hanashite Hon'yaku for Biz



Utsushite Hon'yaku

Tegaki Hon'yaku

Technology Reports　Translation Application with a View to 2020: Tegaki Hon'yaku (P.4)

Screenshot of each translation application

Translation | Character Recognition | Speech Recognition

Technology Reports

Translation Application with a View to 2020: Tegaki Hon'yaku

*NTT DOCOMO has come to provide translation applications for smartphones for use in a variety of scenarios as consumer-oriented translation services. These began with Hanashite Hon'yaku®*1 (speech recognition translation) in 2012 and continued with Mail Hon'yaku®*2 (text translation) and Utsushite Hon'yaku (recognition and translation of camera-captured characters). This article describes NTT DOCOMO's translation services with a focus on the new Tegaki Hon'yaku®*3 (handwriting translation) service.*

Communication Device Development Department

Shin Oguri
Yoshinori Ikehata
Masayoshi Shiota
Misa Tanaka

1. Introduction

Thanks to a weak yen and easing of visa requirements for visitors to Japan as well as an expansion of international airline routes in the Asia region, it appears that the initial target of "20 million foreign tourists by 2020" set by the Japanese government will be reached earlier than expected. New targets have therefore been set in the form of "40 million visitors by 2020 and 60 million visitors by 2030" [1]. Consequently, to provide hospitable services to foreign tourists who are rapidly increasing in number, the need is growing for translation services that can overcome the language barrier and promote communication.

Against the above background, NTT DOCOMO has come to provide smartphone-based translation applications that can be used in a variety of scenarios as translation services for consumers. These began with the release of Hanashite Hon'yaku in 2012, which was followed by the launching of Mail Hon'yaku, Utsushite Hon'yaku, and Jspeak®*4.

There is also "Hanashite Hon'yaku for Biz" that NTT DOCOMO provides as a service for Japanese businesses to support reception of foreign visitors. This article describes NTT DOCOMO's translation services with a focus on a new addition to these services called "Tegaki Hon'yaku" (handwriting translation).

2. Comprehensive View of Translation Applications

2.1 Overview of Translation Applications

NTT DOCOMO's translation applications use four means of input depending on the service: speaking, typing (by keyboard), writing, and camera capture. They also employ a variety of NTT DOCOMO assets*5 (**Figure 1**). Speech recognition and translation technologies are typical of NTT DOCOMO assets used

*1 **Hanashite Hon'yaku®:** A registered trademark of NTT DOCOMO, INC.
*2 **Mail Hon'yaku®:** A registered trademark of NTT DOCOMO, INC.
*3 **Tegaki Hon'yaku®:** A registered trademark of NTT DOCOMO, INC.

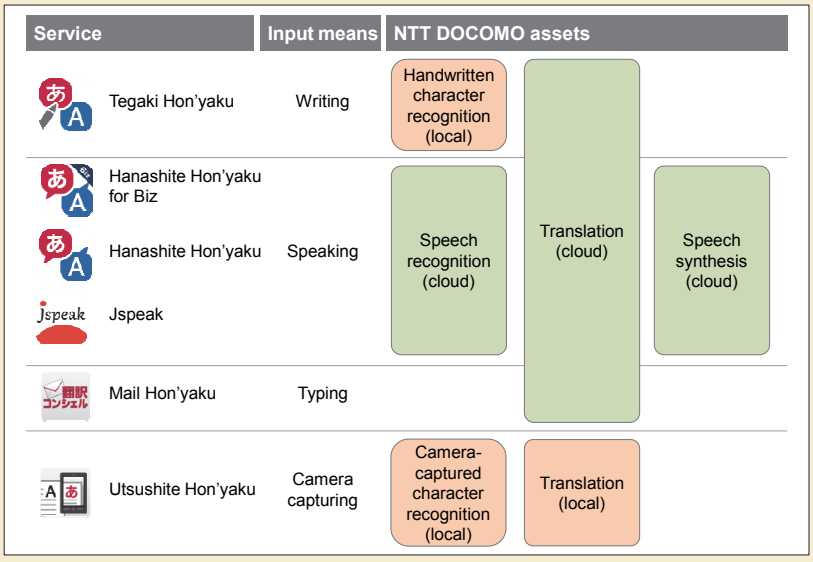


Figure 1 Translation applications and NTT DOCOMO assets

for translation. The recognition and translation of these different forms of input data can be divided into cloud-type and local-type processing. Cloud-type processing provides server-based functions via packet communications and excels in vocabulary size. Local-type processing, on the other hand, provides functions via the smartphone application only, which means no need for packet communications and fast response times. NTT DOCOMO selects either cloud-type or local-type processing for a translation service depending on the features and characteristics of that service. The architecture adopted for implementing these translation services is shown in **Figure 2**. This architecture deploys engines appropriate for each application and language while each service uses whichever engines it needs to function.

2.2 Overview of Each Translation Service

The following presents an overview of each translation service. A screen shot of each translation application is shown in **Figure 3**. A detailed explanation of Tegaki Hon'yaku (handwriting translation) is provided later.

1) For Consumers

(1) Hanashite Hon'yaku

Hanashite Hon'yaku is an application that translates a conversation into each party's language via a smartphone. It is equipped with a "speech translation function" for translating a face-to-face conversation using one smartphone and a "phone-call translation function" for translating a call with a person located elsewhere.

(2) Jspeak

Targeting overseas visitors to Japan, Jspeak provides total support for a wide range of activities related to a Japan visit. The application is equipped with an information-provision function to help visitors learn more about the attractive features of Japan through coupons and information on tourist attractions, popular stores, etc. It also incorporates a communication-support function featuring speech translation and a convenient phrase book. It implements the speech translation function using Hanashite Hon'yaku technology and provides a large collection of phrases.

(3) Mail Hon'yaku

The Mail Hon'yaku application helps to smooth out e-mail exchanges with a person writing in a different language through translation functions for creating and receiving mail.

*4 **Jspeak®:** "Jspeak" and the "Jspeak" logo are registered trademarks of NTT DOCOMO, INC.

*5 **Assets:** Specifically, technology assets.

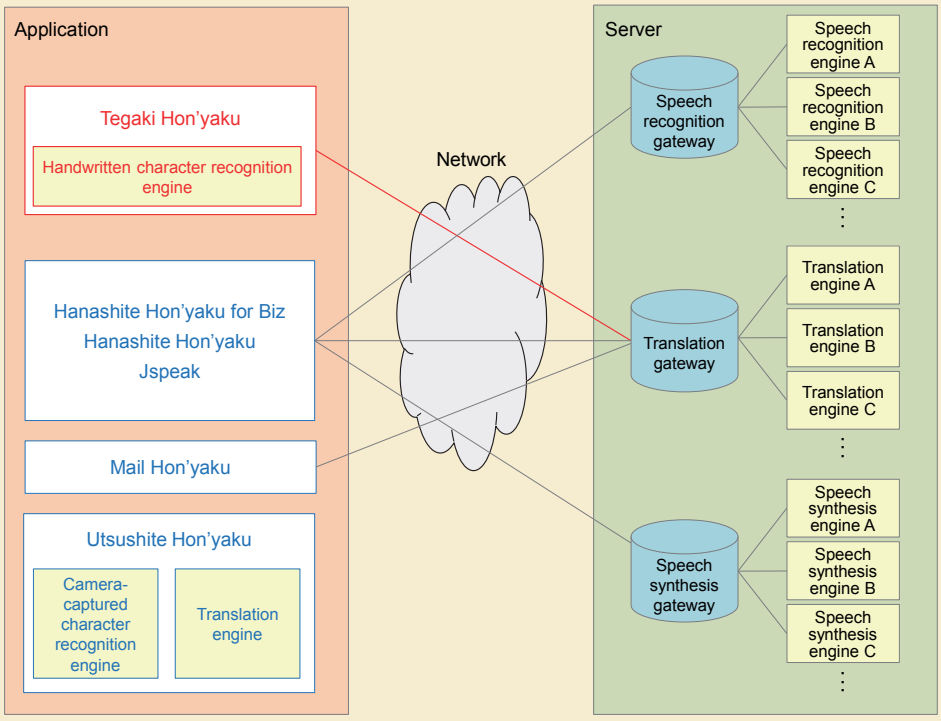
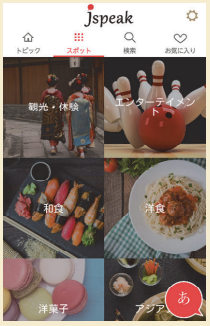


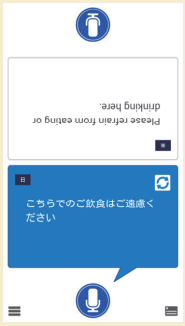
Figure 2 Architecture of translation services

Hanashite Hon'yaku

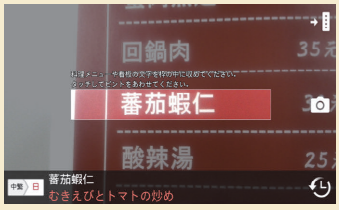


Jspeak

Mail Hon'yaku



Hanashite Hon'yaku for Biz



Utsushite Hon'yaku

Tegaki Hon'yaku

Figure 3 Screenshot of each translation application

Its User Interface (UI) assumes linking with other applications and features enhanced linking with docomo mail in particular.

(4) Utsushite Hon'yaku

This application provides real-time display of translation results when a picture taken by the smartphone's camera includes characters representing a place name, an item from a restaurant menu, etc. Translation is performed by a camera-captured character recognition engine and translation engine incorporated in the application itself. This enables the user to use a translation function without having to use packet communications.

2) For Business (Hanashite Hon'yaku for Biz)

Hanashite Hon'yaku for Biz targets airports and railroad companies in the transportation industry and hotels and inns in the lodging industry as well as restaurants, retail establishments, etc. as an application supporting the reception of foreign visitors to Japan [2].

This application integrates a speech translation function based on consumer-oriented Hanashite Hon'yaku technology, a collection of phrases having a high frequency of use in reception settings, and a video calling function for connecting with a bilingual operator to deal with more complicated conversations.

Trials were performed repeatedly with companies in various industries to make the application easy to use by receptionists and easy to understand by customers. As a result, the application was designed to support not only over-the-counter reception scenarios but also joint viewing of a screen by the receptionist and customer. Hanashite Hon'yaku for Biz supports ten languages: English, Chinese, Korean, Italian, French, Portuguese, German, Spanish, Thai, and Indonesian.

NTT DOCOMO is putting a great deal of effort in spoken-language translation assuming diverse usage scenarios including Hanashite Hon'yaku for Biz. It is in the process of developing technologies and services applicable to meeting minutes and SNS postings where translation needs are high [3].

3. Tegaki Hon'yaku

3.1 Overview of Tegaki Hon'yaku

Tegaki Hon'yaku is an application for translating handwritten text. It currently supports six languages: English, Chinese (Mandarin), Chinese (Taiwanese), Korean, French, and Spanish. It achieves recognition of handwritten text by a built-in recognition engine that analyzes the locus and speed of a pen tip moving on the screen and converts the results of that analysis to character data.

The application is also equipped with a drawing function as a complement to the handwriting translation function. This function can be used, for example, to draw a map on a smartphone screen when giving directions to a foreign visitor. If the user then handwrites words on top of that map, the application will translate and display those words in the visitor's language in real time.

An example of such a text conversion is shown in **Figure 4**. This example shows Japanese words written on a map drawn with the drawing function and the translation of those words into English.

In this way, it becomes easy to communicate something that may be difficult to convey by words only.

3.2 Background to Development of Tegaki Hon'yaku

1) Issues with Hanashite Hon'yaku

The following two issues came to light when attempting to analyze a customer's voice with the previously launched Hanashite Hon'yaku application.

(1) Speech translation could not be used in a noisy environment that made conversation difficult.

(2) Asking foreign visitors with no experience in using a translation service to talk into a smartphone or tablet was a high hurdle to overcome.

2) Solutions

Tegaki Hon'yaku was developed to resolve these issues using the following two methods.

- For issue (1): Change the input method from speech to handwriting to eliminate the effects

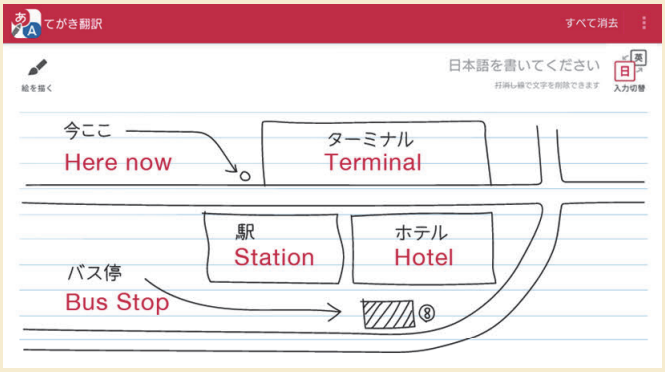


Figure 4 Example of handwriting translation

of noisy environments.

- For issue (2): Make use of everyone's experience of communicating by handwritten text to make it easy to use a translation application even for users resistant to talking into a smartphone.

During initial studies for this application, we were particularly aware of the need for a UI that would preserve the natural act of "handwriting" and support communication in the form of writing characters on a guidebook or map to ask or give directions. Furthermore, in the trial period, we wanted even customers with no opportunity to take an overseas trip to experience Tegaki Hon'yaku. We therefore asked subjects to try a version of the application equipped with a function for translating handwritten text into a regional dialect and received a good evaluation (the commercial version does not support such a dialect function).

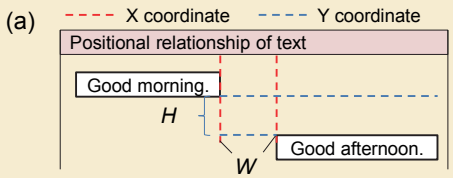
3.3 Translation-range Judging Logic

The initial version of the application had no function for automatically judging the translation range of sentences written on the screen. As a result, translation errors would sometimes occur when continuing a sentence that could not fit on the screen onto another line or when writing two sentences on the same line. We therefore equipped the application with translation-range judging logic that uses the coordinates of handwritten text to decide whether to translate that text as one sentence or multiple sentences (**Figure 5**). The process flow of this logic is as follows.

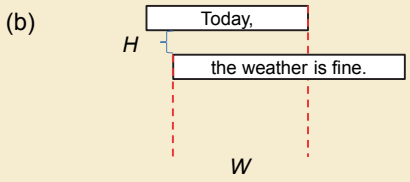
(1) Determine whether handwritten text is on one line or multiple lines.

(2) If on one line, compare each character with the one to its immediate right in order from left to right and use the interval between two such characters to judge whether the handwritten text consists of one sentence or multiple sentences. Here, H in Fig. 5 (a) is the distance between the bottom edge and top edge of adjacent characters, while W is the distance between the right end and left end of adjacent character strings.

(3) If on multiple lines, compare the coordinates of characters on the first line with those on the second line to judge whether the handwritten text consists of one sentence or multiple sentences. Here, H in Fig. 5 (b) is the distance between the bottom edge of the characters on the first line and the top edge of the characters on the second line, while W is the distance between the right end of the character string on the first line and the left end of the character string on the second line.

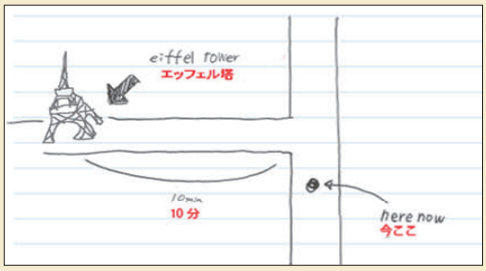


Judge as multiple sentences if *W* and *H* are both greater than certain values.



Judge as one sentence if *W* and *H* are both within certain values.

Example of multiple sentences



Example of one sentence

Figure 5 Judging translation range

4. Tegaki Hon'yaku Trials

4.1 Tokyo International Air Terminal Corporation

1) Experiment Overview

In cooperation with Tokyo International Air Terminal Corporation, NTT DOCOMO introduced Tegaki Hon'yaku at information counters manned by guides and concierges inside the Haneda Airport International Passenger Terminal and launched a trial in September 2015 with the aim of enhancing the assistance given to foreign visitors [4].

In this trial, Tegaki Hon'yaku was used to help foreign visitors who had just arrived in Japan and were unsure of how to proceed. For example, it could be used to draw maps and give directions to where one would be staying or information on airport facilities. The trial also included interviews with guides and concierges to hear about their impressions of using the Tegaki Hon'yaku application.

2) Issues Uncovered by Experiment

The results of the experiment showed that Tegaki Hon'yaku could be an effective means of communication even in noisy environments or places where verbal communication is difficult. However, they also revealed that the time taken up by handwriting on a screen

and translating that input could sometimes be unacceptably long relative to the reception speed required for handing many inquiries at an information counter.

3) Improvement Measures

To resolve the above issues, the application was equipped with a "point-to-it phrase book," which facilitates communication by displaying frequently used phrases that can be pointed at to convey intent, a "suggest function" that can call up fixed phrases from a single handwritten word, and a function for linking to Hanashite Hon'yaku for Biz that can provide machine translation of speech and access to and interpretation by a human operator.

These functions can be used as needed by a receptionist according to the current situation. For example, the "point-to-it phrase book" can be used to provide a quick response when dealing with a foreign visitor for the first time, "writing" by Tegaki Hon'yaku can be used when solving a problem by the "point-to-it phrase book" is difficult, and "speaking" by Hanashite Hon'yaku for Biz can be used to enable verbal communication (**Figure 6**).

We also added a screen rotation function to facilitate communication and changed to a UI that can smoothly request re-input when the intention of the sentence to be translated cannot be understood.

4.2 Tokio Marine & Nichido Fire Insurance Co., Ltd.

A trial use of Tegaki Hon'yaku was performed at the Kaijo Building Clinic of the Tokio Marine & Nichido Fire Insurance Co., Ltd to receive visiting and resident foreigners and process their applications for medical interviews [5].

In this trial, we assessed the application in conjunction with business needs of the Kaijo Building Clinic. Specifically, we evaluated the translation performance of terms and sentences actually used in clinic settings and used the results obtained for developing the application further.

5. Conclusion

This article described NTT DOCOMO translation applications with a focus on Tegaki Hon'yaku (handwriting translation). With a view to 2020, NTT DOCOMO plans to improve the accuracy of these applications and increase the number of supported languages with the aim of providing translation services of even

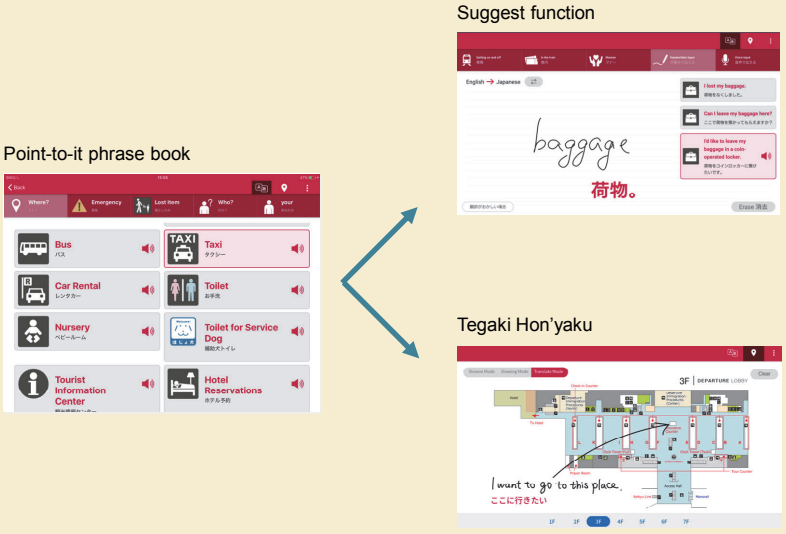


Figure 6 Screen shots of trial application used with Tokyo International Air Terminal Corporation

higher value for its customers.

References

[1] Prime Minister of Japan and His Cabinet: "Summary of Proceedings of 2nd Meeting of the Council for the Development of a Tourism Vision to Support the Future of Japan," Mar. 2016. http://japan.kantei.go.jp/97_abe/actions/201603/30article1.html

[2] NTT DOCOMO Press Release: "NTT DOCOMO Launches Hanashite Hon'yaku for Biz—Customer-interaction Translation Service for Business," Jun. 2016 (in Japanese). https://www.nttdocomo.co.jp/info/news_release/2016/06/01_00.html

[3] M. Takeichi et al.: "Spoken Language Translation Services for Diverse Usage Scenes," NTT DOCOMO Technical Journal, Vol.18, No.3, pp.12-22, Jan. 2017.

[4] NTT DOCOMO Press Release: "DOCOMO Launches Trial of Handwriting Translation Service Tegaki Hon'yaku," Sep. 2015. https://www.nttdocomo.co.jp/english/info/media_center/pr/2015/0909_00.html

[5] Tokio Marine & Nichido Fire Insurance Co., Ltd.: "New Products and Services for Foreign Visitors to Japan and Inbound Businesspersons," May 2016 (in Japanese). http://www.tokiomarine-nichido.co.jp/company/release/pdf/160519_01.pdf

Machine Translation | Spoken Language Translation | Preprocessing Technology

Technology Reports

Spoken Language Translation Services for Diverse Usage Scenes

The number of foreign visitors to Japan continues to increase with expectations of 40 million visitors by 2020. With such a large number of foreign guests coming to Japan, there are hopes for an environment that can provide stress-free communication. At the same time, the globalization of business is accelerating resulting in even more situations where multilingual communication is needed.

Among the translation services that NTT DOCOMO has decided to develop, this article describes the issues and solutions surrounding meeting translation, SNS translation, and customer-reception translation as services that translate spoken language for diverse usage scenes.

Service Innovation Department

Masato Takeichi
Takaya Ono
Yuki Chijiwa
Yixin Jiang

1. Introduction

The number of foreign visitors to Japan in 2015 reached 19,740,000, which surpassed the previous record set the year before in 2014 [1]. With this in mind, the Japanese government has set 40 million visitors as the target for 2020 with expectations that an environment enabling all foreign visitors to communicate freely can be provided [2]. At the same time, the number of employees of overseas subsidiaries of Japanese firms increased to 5,750,000 in FY2014 from 4,990,000 in FY2010 [3], which reflects the onward march of corporate globalization in which multilingual communication will be increasingly needed.

With a view to 2020, NTT DOCOMO has undertaken the development of speech recognition*1 technology and machine translation*2 technology and services that apply those technologies toward means of communication that can surmount language and cultural barriers.

In **Figure 1**, translation services are classified along two axes representing "spoken language – written language" and "hard/soft textual styles of writing" to clarify the technical issues that must be addressed. NTT DOCOMO sets the upper-right area of the figure—"spoken language" and "soft"—as a near-term target with a focus on foreign visitors to Japan and aims to improve the accuracy of speech recognition and machine translation to meet this target. The following three services are currently being developed as part of this endeavor.

(1) Meeting translation: Targeting meetings in different languages, this service translates meeting speech into the user's native language for verbal readout and

*1 **Speech recognition:** Technology for converting speech signals generated by human utterances into text.

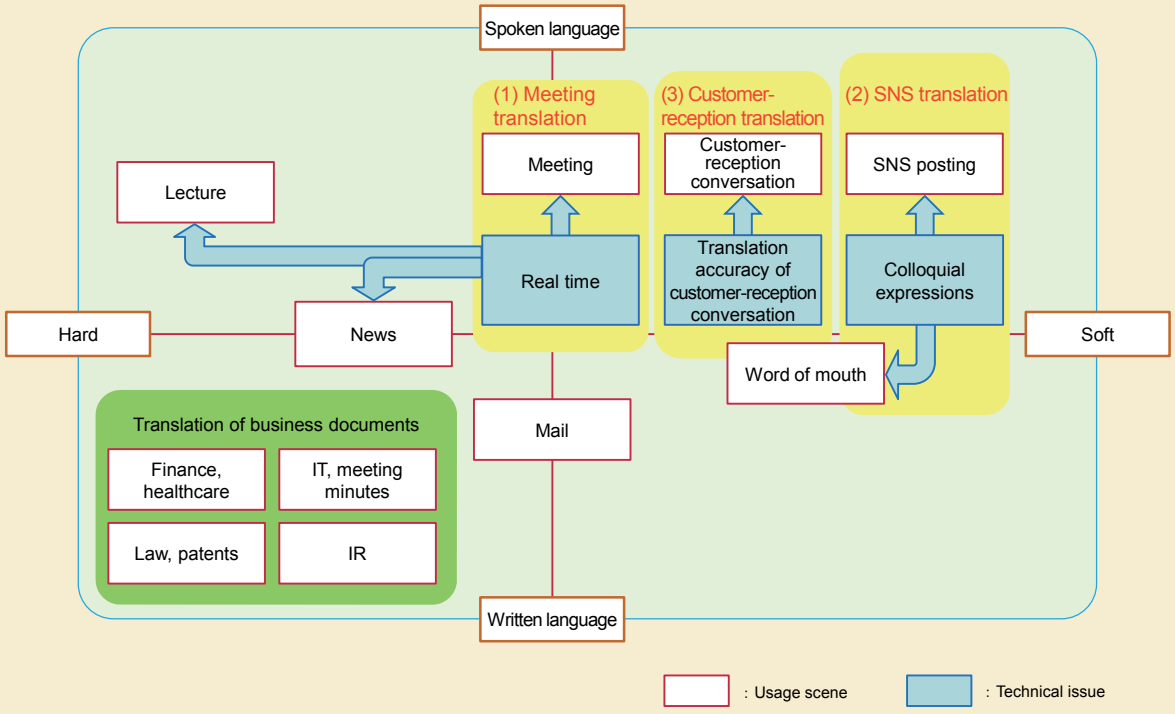


Figure 1 Usage scenes of translation services and technical issues

text display.

(2) SNS translation: This service provides a text translation of SNS posts that are normally full of colloquial expressions. Although it may appear at first glance that such posts consist of written language, they actually include many expressions that can be classified as "soft" or "spoken language."

(3) Customer-reception translation: This service performs speech translation of customer-reception conversation between a customer and staff at a store or other business site. For simple communication needs, it provides an easy means of waiting on a customer through speech recognition and machine translation, and for complex conversation needs, it provides a means of talking with a remote interpreter. Both of these means combined enable accurate and stress-free communication. Since the development of a prototype in 2014, NTT DOCOMO has performed trials with a number of companies and has repeatedly improved the User Interface (UI)*3 and raised the accuracy of speech recognition and machine translation. A commercial service was finally launched in June 2016 as "Hanashite Hon'yaku for Biz"*4.

In this article, we provide an overview of these three services, introduce issues associated with each of these services, and describe NTT DOCOMO's approach to resolving those issues.

2. Overview of Each Translation Service and Associated Issues

2.1 Overview of Translation Services

A translation service normally combines three technologies: speech recognition, machine translation, and speech

*2 **Machine translation:** Technology for mechanically converting sentences and words in a certain language into another language and outputting the results. There are two main methods: rule-based machine translation and statistical machine translation.

*3 **UI:** Operation screen and operation method for exchanging information between the user and computer.

*4 **Hanashite Hon'yaku for Biz:** An NTT DOCOMO business-oriented translation service supporting customer reception by Japanese store clerks for foreign visitors to Japan. Launched in June 2016. A registered trademark of NTT DOCOMO, INC.

synthesis*5. The system configuration of such a service is shown in **Figure 2**.

(1) Speech recognition technology is used to convert speech to text. This process first identifies an utterance by utterance-interval detection*6 where the end of an utterance is established by a silent interval*7 lasting longer than a certain period of time. It then removes noise in the background by noise removal*8 technology. Finally, it converts the results of speech recognition to text using an acoustic model*9 and language model*10 in the speech recognition engine.

(2) Machine translation technology is used to mechanically translate input text into text in another language. This process begins with a preprocessing step that performs syntactic analysis*11 and named entity classification*12 against the pre-translated text to improve translation accuracy. It then uses a machine translation engine*13 and its translation model*14 and language model to translate the input text into text of the designated language. Finally, as a post-processing step, it uses bilingual data from a dictionary to perform substitution of previously classified words before preparing and outputting the results of translation.

(3) Speech synthesis technology produces artificial speech data from input text (results of translation) and verbally reads out that data.

The total corpus of a translation service consists of a corpus for speech recognition and a bilingual corpus for machine translation. The former consists of sets of uttered speech and text created by logs accumulated during service provision. The latter consists of bilingual sets of text in the pre-translation language and text in the post-translation language in units of sentences (e.g., "いらっしゃいませ" ⇔ May I help you?).

To improve the accuracy of speech

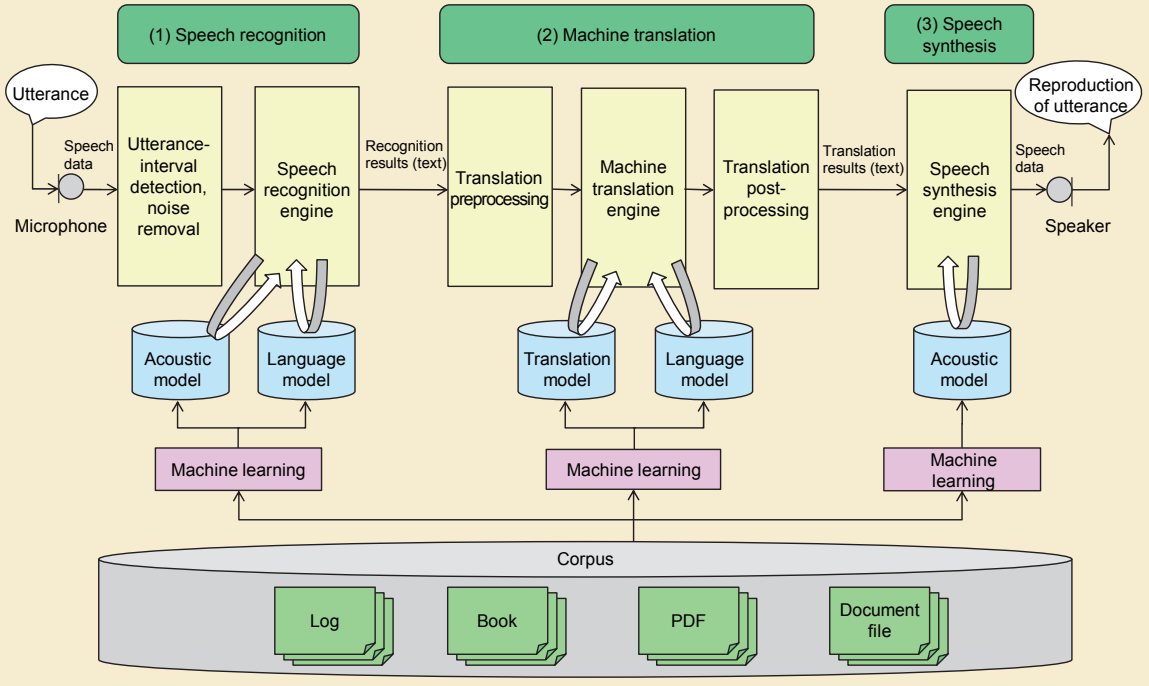


Figure 2 System configuration of speech translation technologies

*5 **Speech synthesis:** Technology for artificially creating speech data from text and verbally reading out text.

*6 **Utterance-interval detection:** Technology for determining the intervals in speech signals in which speech exists and those in which it does not exist.

*7 **Silent interval:** Interval determined to be absent of speech.

*8 **Noise removal:** Process of removing street noise, utterances of other people, etc. to recognize the speech of a specific person.

*9 **Acoustic model:** Statistical model comprising frequency characteristics of phonemes targeted for recognition possesses.

*10 **Language model:** Statistical model comprising morpheme arrangements and frequency of those arrangements.

*11 **Syntactic analysis:** Technology for analyzing the structure of inter-phrase dependency and clarifying grammatical relationships in a sentence.

recognition and machine translation, it is generally necessary to collect a large volume of sentences for each language pair*15 (e.g., Japanese/English, Japanese/Chinese, Japanese/Korean, etc.) as a bilingual corpus of frequently used sentences in usage scenes marked by the red frames in Fig. 1 (such as meetings, SNS posts, and customer reception). Such a large bilingual corpus can then be used as training data in machine learning*16 so that translation models and language models can be created specific to usage scenes. Likewise, in speech recognition, machine learning can be used to create acoustic models and language models using corpuses corresponding to different usage scenes as learning data. In speech synthesis, too, machine learning can be used to create acoustic models.

2.2 Issues Associated with Each Translation Service

The issues particular to each translation service are described below.

(1) Meeting translation must be able to display the results of speech recognition and machine translation in real time and correctly detect the breaks between utterances. With prior technology, the results of applying speech recognition to the utterances of a meeting participant would be displayed in text after an utterance was completed, which would give the user a strong sense of having to wait until results were displayed.

(2) A major issue in SNS translation is dealing with colloquial expressions that frequently appear in SNS posts. Prior technology was incapable of dissecting pre-translated Japanese text written in a non-standard, colloquial format such as "おたんじょーびおめでとぉー" and "これはヤヴァい."

(3) A customer-reception translation service must achieve courteous interaction with customers by store staff and deal with colloquial questions from customers. However, there is a lack of bilingual corpuses for conversation that occurs when waiting on customers, so the issue here is finding ways of improving the accuracy of machine translation for dialog spoken in customer-reception scenarios.

3. Meeting Translation

As shown in **Figure 3**, the meeting translation service performs speech recognition in real time of the conversation of meeting participants speaking different languages and translates that conversation into each other's language. It is also capable of text translation by keyboard input so that a user can participate in a meeting by text if speaking out loud is not possible for whatever reason.

1) Improvement of Real-time Characteristics

To achieve a meeting with good tempo, it is desirable that the processing delay experienced by meeting participants for both speech recognition and machine translation be made small to provide good real-time characteristics.

For this reason, the speech recognition process sequentially displays the results of maximum likelihood estimation for each morpheme*17. This has the effect of minimizing the delay between the utterances of meeting participants and the display of those utterances thereby improving the real-time characteristics experienced by the participants.

2) Automatic Detection of Utterance Interval

Some services that employ speech recognition have the user press a button to mark the beginning and end of an utterance to unambiguously indicate the user's utterance interval. However, for a meeting translation service in which utterances are continuously and rapidly exchanged among meeting participants, a more desirable approach is to automatically detect a speaker's utterance interval and forgo the use of start/stop buttons.

In addition, utterances may be delimited by the continuance of a silent interval that occurs while the speaker is thinking about the next thing to say or

---

*12 **Named entity classification:** Technology for replacing named entities in the input sentence with labels expressing proper nouns before machine translation and for replacing those labels using a dictionary after machine translation.

*13 **Machine translation engine:** Software for statistically translating text using a language model and translation model trained for machine translation.

*14 **Translation model:** Statistical model used for calculating the extent to which words in a pair of sentences in the pre-translation language and post-translation language semantically correspond with each other.

*15 **Language pair:** A combination of two languages as the translation source and translation objective (e.g. English/Japanese).

*16 **Machine learning:** A mechanism allowing a computer to learn the relationship between inputs and outputs, through statistical processing of example data.

*17 **Morpheme:** Smallest meaningful unit of a language.

simply hesitating. Specifically, in the case of **Figure 4** (1), the Japanese spoken in this example was divided into two utterances consisting of "このアプリは音声認識の結果を" and "リアルタイムに表示します." As a consequence, the machine translation also output two utterances consisting of "This application is speech recognition results." and "Real-time display." In other words, the results output by machine translation in this case differed from the speaker's intention.

To solve this problem, we adjusted the parameter that judges the silent interval so that input speech could be detected as a continuous utterance even with the occurrence of some hesitation

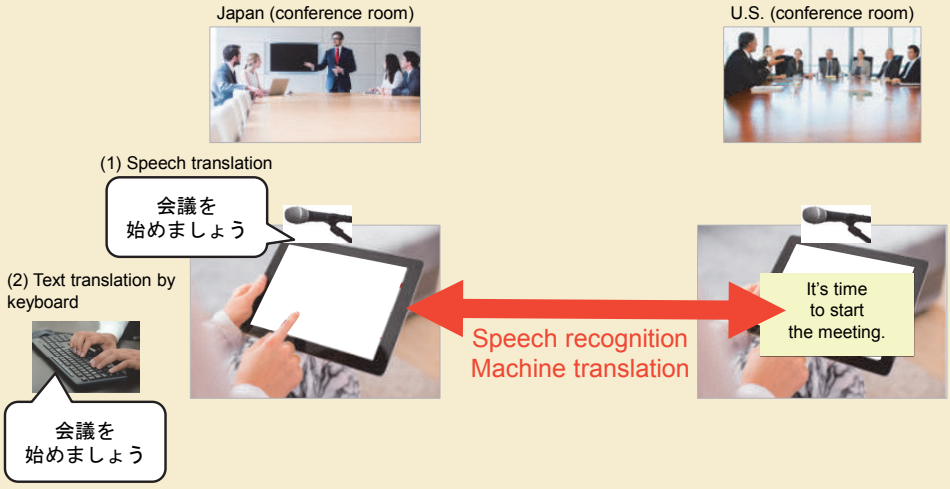
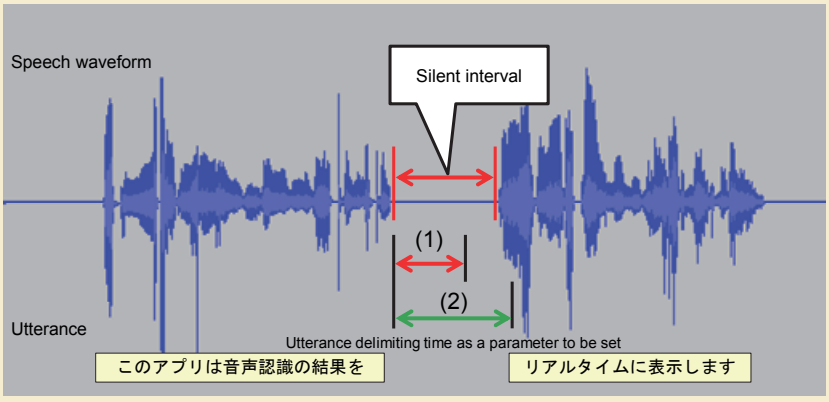


Figure 3 Overview of meeting translation service



Translation results

Case (1)
「このアプリは音声認識の結果を」→ "This application is speech recognition results."
「リアルタイムに表示します」→ "Real-time display."
Case (2)
「このアプリは音声認識の結果をリアルタイムに表示します」
→ "This application is displayed in real-time voice recognition results."

Figure 4 Silent interval caused by hesitation

while speaking. We did this by extracting the silent intervals that occur when speakers hesitate while speaking in meetings and calculating their average value. We then used this average value as a basis for optimizing the parameter to be set as the time for delimiting utterances and performed a test to see if utterances in a meeting could be correctly delimited.

As shown in Fig. 4 (2), adjusting the parameter in this way prevented the silent interval from being treated as an utterance delimiter resulting in the single Japanese utterance "このアプリは音声認識の結果をリアルタイムに表示します." The correct machine translation could then be output as "This application is displayed in real-time voice recognition results."

4. SNS Translation

SNS translation is a service that can translate colloquial sentences and expressions in social media to other languages. The process flow of the SNS translation system is shown in **Figure 5**. This system differs from the other translation services introduced here in that it consists only of the machine translation section shown in Fig. 2.

However, the system adds two functions to translation preprocessing in the usual machine translation section. The first one is sentence dissection and the second one is normalization of variant character strings.

- In addition to identifying ordinary sentence delimiters, the sentence dissection function determines the presence of parentheses, emoticons (e.g. ヽ(・∀・)ﾉ ｷｬｯ), service-unique symbols (e.g. RT), URLs, and onomatopoeic/mimetic words in a sentence. As a result, only that part of the sentence recognized as text for translation can now be passed to the following normalization function for processing.
- The normalization function enhances the morphological analysis of traditionally written, standard Japanese by dividing the sentence into "tokens" using variant-token morphological analysis [4] [5] that can deal with colloquial expressions. Here, "tokens" are the result of dividing the input text into minimal word units of the

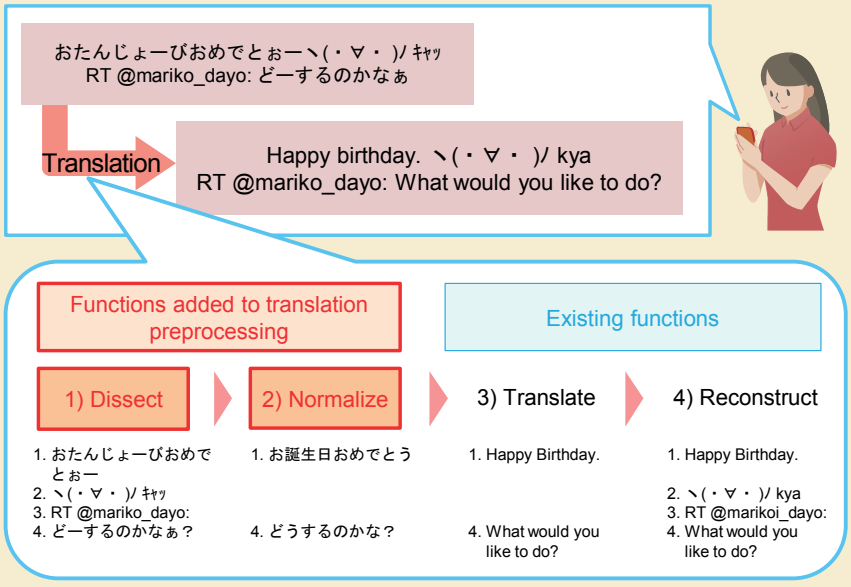


Figure 5 Process flow of SNS translation system

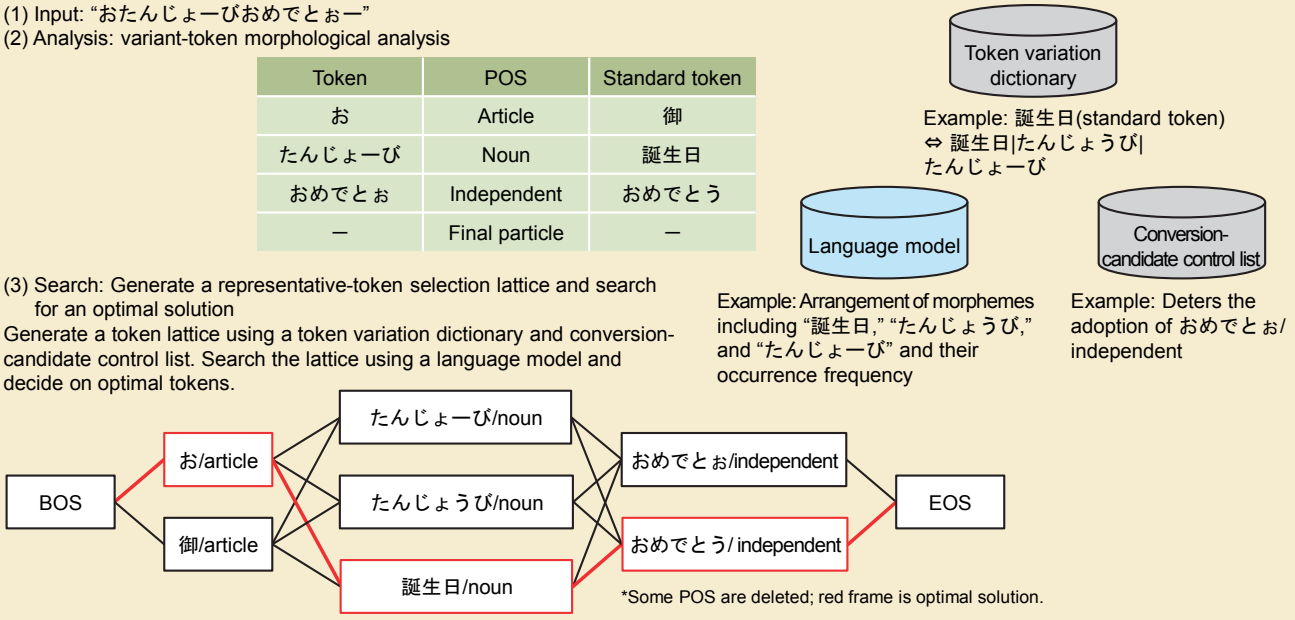
Japanese language. Next, the function generates a representative-token selection lattice*18 using a token variation dictionary and conversion-candidate control list. It then converts the variant token into an optimal token using a language model based on the huge corpus used in machine learning by the machine translation engine. This function can also deal with ambiguity by substitution processing based on negative/positive determination.

1) Generation of Representative-token Selection Lattice and Conversion of Variant Tokens

We here explain the mechanism of generating a representative-token selection lattice and converting variant tokens using **Figure 6**.

First, referring to Fig. 6 (1) and (2), the results of variant-token morphological analysis are used to prepare tabular data consisting of token, Parts Of Speech (POS), and standard token*19 entries from the input sentence. In this example, the input sentence "おたんじょーびおめでとぉー" is divided into the tokens "お," "たんじょーび," "おめでとぉ," and "ー," and the standard tokens "御," "誕生日," "おめでとう," and "ー" are given for each of these. Here, "standard token" refers to standard notation as used in newspapers and other publications.

Next, focusing on standard tokens, a lattice of selectable tokens is created using tokens in the conversion-candidate control list. This list includes candidates not desirable for conversion and candidates uniquely desirable for conversion



Token	POS	Standard token
お	Article	御
たんじょーび	Noun	誕生日
おめでとぉ	Independent	おめでとう
ー	Final particle	ー

Figure 6 Mechanism of variant-token morphological analysis and variant-token conversion

*18 **Lattice:** A lattice-shaped collection of data that arranges a series of morphemes in the horizontal direction and morphemes with the same meaning but different notation in the vertical direction.

*19 **Standard token:** A characteristic string expressed in characters and symbols that are written in standard format according to undeviating grammar.

as tokens linked to the standard token but not included in the input sentence. For example, tokens linked to the standard token "誕生日" include "たんじょーび," "たんじょうび," and "誕生日," all of which are used to create the token lattice shown in Fig. 6. Here, "token lattice" refers to a graphical structure that enumerates selectable tokens. Now, using a language model employed by the machine translation engine, the strings of tokens in this token lattice are searched to determine an optimal sequence of tokens, as shown in Fig. 6 (3). In this example, the optimal solution is found to be the tokens "お," "誕生日," and "おめでとう," as shown in Fig. 6 (4). This representative-token conversion process results in the normalization of variant character strings.

Incidentally, based on technical support and process results provided by NTT Media Intelligence Laboratories, we have incorporated the language models and conversion-candidate control list created by NTT DOCOMO in this variant-token morphological analysis function and representative-token conversion function and have implemented these functions in the translation preprocessing section.

2) Negative/positive Determination

Japanese can sometimes be ambiguous making it difficult to determine intention from one word or even one sentence. For example, a SNS post stating "これはヤヴァい (Kore wa yabai!)" in relation to a food review can be interpreted as either "delicious" or "bad tasting" if translating one sentence at a time. Accordingly, by applying negative/positive determination*20 to previous and subsequent sentences, the above expression can be substituted by "This is great!" for a positive sentiment and "This is terrible!" for a negative sentiment. Such substitution processing can deal effectively with ambiguity.

In SNS translation, machine translation is applied only to those portions of the input sentence that have been normalized by the above sentence dissection function and variant-character-string normalization function. In translation post-processing, the structure of the input sentence is reconstructed integrating the results obtained in the machine translation step and the reconstructed sentence is returned to the user as the translated sentence.

As a result of this translation post-processing, the accuracy of translating "soft" spoken language from Japanese to English, Chinese, and Korean was found to outperform typical engines of other companies. The results of comparing the accuracy of Japanese/English translation between NTT DOCOMO's SNS translation engine and another company's engine are shown in **Figure 7**. For this comparison, we extracted 100 sentences randomly from SNS posts related to food, cosmetics, and travel and performed a subjective evaluation using three subjects based on the following criteria: "○ (3 points): correctly translated" and "△ (2 points): some errors noticed but meaning is conveyed." The average of the scores given by the three subjects was calculated for each sentence. On tabulating the total score for all sentences, it was found that NTT DOCOMO's SNS translation engine outperformed the other company's engine by 45 points.

5. Customer-reception Translation

Customer-reception translation requires high accuracy in machine translation for both the case of translating courteous sentences spoken by a Japanese-speaking clerk into multiple languages and the case of translating questions that may not necessarily be formal or courteous in the customer's foreign language into Japanese. With this in mind, we created a translation model especially for customer reception by collecting sentences frequently spoken when waiting on customers as a bilingual corpus and having a translation engine learn those sentences. We also applied this bilingual corpus to the language model of the speech recognition engine so that it too could deal with conversation in customer reception.

The following describes the process flow for compiling the bilingual corpus.

*20 **Negative/positive determination:** Method for determining whether the writer's intention is negative or positive from that document.

1) Collection of Common Bilingual Corpuses

The flow of collecting commonly used bilingual corpuses is shown in **Figure 8**. The first step in this process is to collect bilingual data as the basis of a bilingual corpus by the following methods:

- Transcribe and collect logs of commercial services (Hanashite Hon'yaku, mail translation, etc.) in conformance with terms of service
- Purchase and use external bilingual corpuses

The next step is to format the bilingual data collected by these two methods

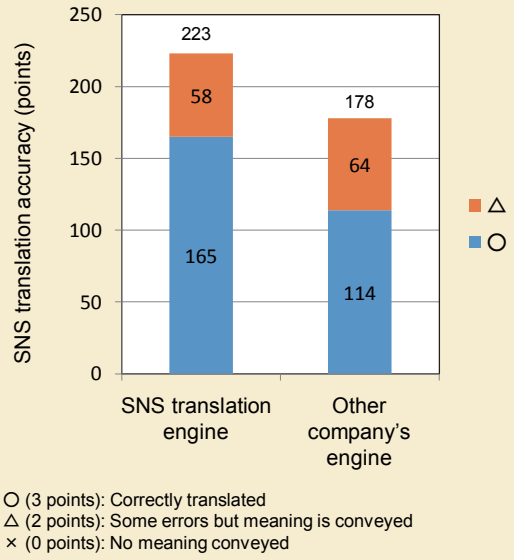


Figure 7 Subjective evaluation of SNS translation accuracy (J-E translation)

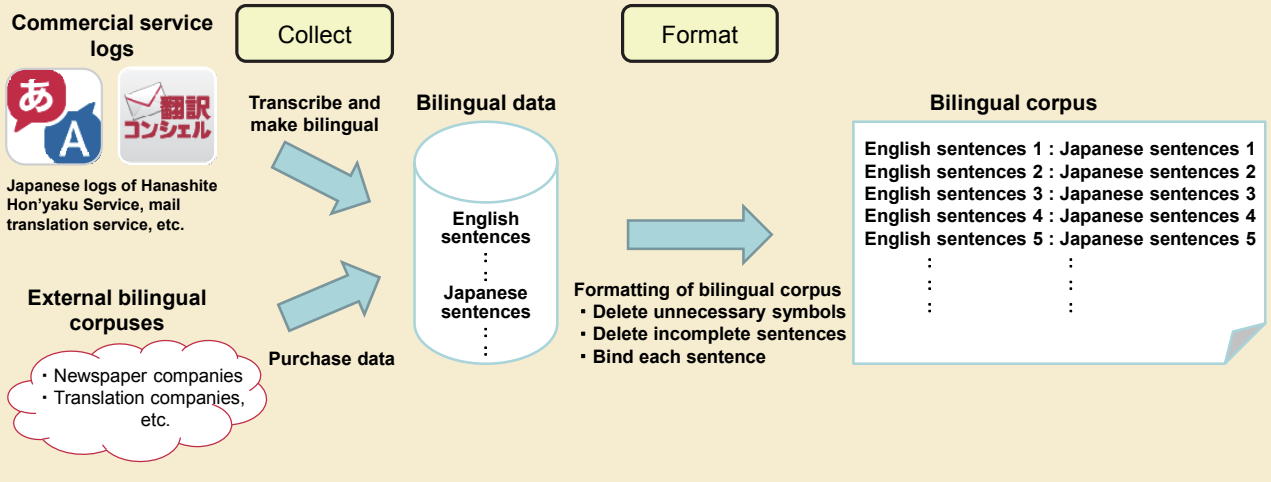


Figure 8 Compilation of bilingual corpus

in the following ways to create a bilingual corpus:

- Delete unnecessary symbols included in the collected corpuses
- Delete incomplete sentences
- Bind each sentence

2) Create a Bilingual Corpus Manually

This time, with the aim of increasing corpus content specialized for a certain usage scene in addition to using the usual collection methods described above, we employed crowdsourcing*21 and other techniques to manually create a bilingual corpus assuming customer-reception conversation. Using the bilingual corpus obtained in this way, we created a translation model and a speech-recognition language model for use in customer reception.

As a result, the accuracy of speech recognition and machine translation was improved for a variety of languages. The results of comparing the accuracy of Japanese/English translation among various engines are shown in **Figure 9**. For this comparison, we randomly extracted 200 sentences from documents and logs related to customer reception and performed a subjected evaluation using five subjects. As shown, the speech-recognition/machine-translation engine that we prepared for customer reception achieved an average value of 218.4 points, which outperformed typical engines of other companies.

6. Conclusion

In this article, we described NTT DOCOMO's approach to achieving spoken language translation and solving associated technical issues with a focus on meeting translation, SNS translation, and customer-reception translation.

Going forward, we plan to refine this technology to make it even easier to use. We also plan to research and develop technologies for improving translation accuracy even further. These will include "prereordering technology" for reordering the words in the target sen-

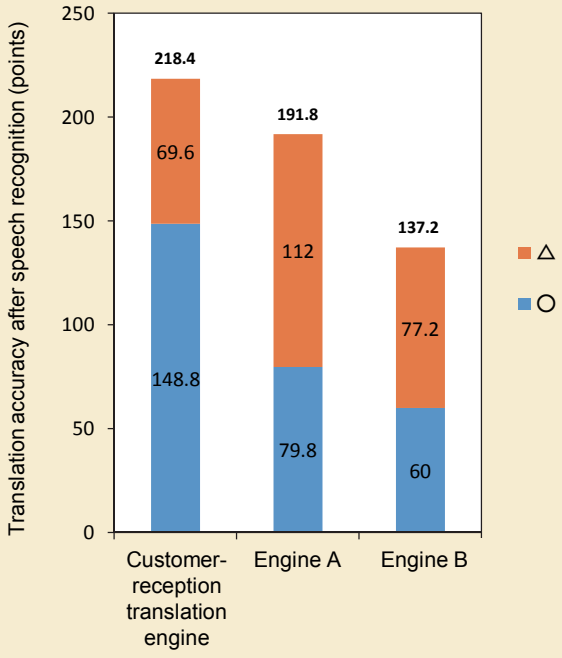


Figure 9 Subjective evaluation of customer-reception translation accuracy (J-E translation)

*21 **Crowdsourcing:** A coined term combining "crowd" and "outsourcing" referring to a new employment format that consigns work or tasks among a number of widely distributed people.

tence prior to translation to mitigate drops in accuracy caused by differences in word order between the source language and destination language, and natural language processing technology for filling in abbreviations and omissions that frequently occur in spoken language.

References

[1] Japan Tourism Agency: "White Paper on Tourism in Japan (2016)."

[2] Prime Minister of Japan and His Cabinet: "Meeting of the Council for the Development of a Tourism Vision to Support the Future of Japan," Mar. 2016. http://japan.kantei.go.jp/97_abe/actions/201603/30article1.html

[3] Ministry of Economy, Trade and Industry, Minister's Secretariat, Research and Statistics Department, Structural Statistics Office, Trade and Economic Cooperation Bureau, Trade and Investment Facilitation Division: "Summary of the 45th Basic Survey on Overseas Business Activities—FY2014 Results—," Jul. 2015. http://www.meti.go.jp/english/statistics/tyo/kaigaizi/pdf/h2c408je.pdf

[4] I. Saito, K. Sadamitsu, H. Asano and Y. Matsuo: "Japanese Morphological Analysis Based on Alignment of Standard and Variant Character Strings and Variant Token Normalization using Character-type Conversion," Proc. of the 20th Annual Meeting of The Association for Natural Language Processing, pp.777-780, Mar. 2014 (in Japanese).

[5] I. Saito, K. Sadamitsu, H. Asano and Y. Matsuo: "Token Normalization and Morphological Analysis using Occurrence Probability of Variant Tokens," Proc. of the 21th Annual Meeting of The Association for Natural Language Processing, pp.51-54, Mar. 2015 (in Japanese).

Weathercast Linking | Sunlight Control Mode | Autonomous/Regeneration Mode

Technology Reports

Weathercast-linked Control in Environmentally Friendly, Disaster-resistant Green Base Stations and Field Trials

Recent years have seen a full-scale rollout of green base stations applying environmentally friendly equipment such as solar panels and lithium-ion batteries to provide an environmental contribution and a countermeasure to disasters. This article describes "weathercast-linked control" as a technology for improving the functional performance of green base stations while achieving the contradictory objectives of an environmental contribution and disaster countermeasure. The technology can be used to control lithium-ion batteries based on information in weather reports, reduce the use of commercial power, and extend service time during a power outage.

Research Laboratories
Masaki Nakamura
Kazuaki Kimura
Kazuhiko Takeno

1. Introduction

In the wake of recent power outages brought on by natural disasters and other events, enhancing power-backup measures for mobile communications is becoming increasingly important. Although radio base stations are already equipped with storage batteries for backup purposes during power outages in the utility grid, it is thought that this scheme can be further enhanced by changing from lead-acid batteries to Lithium-ion Batteries (LiB) to increase full-charge battery capacity [1] [2].

At the same time, studies are being performed on migrating to natural energy as an environmental contribution. In this regard, particular attention is being given to photovoltaic generation, which can generate power during daylight hours when demand is high and therefore reduce demand for commercial power during peak hours. Looking forward, the introduction of natural energy such as photovoltaic power in radio base stations can also be envisioned, and it is said that "green base stations" that can effectively combine storage batteries with natural energy and reduce environmental load can be an effective means of doing so [3].

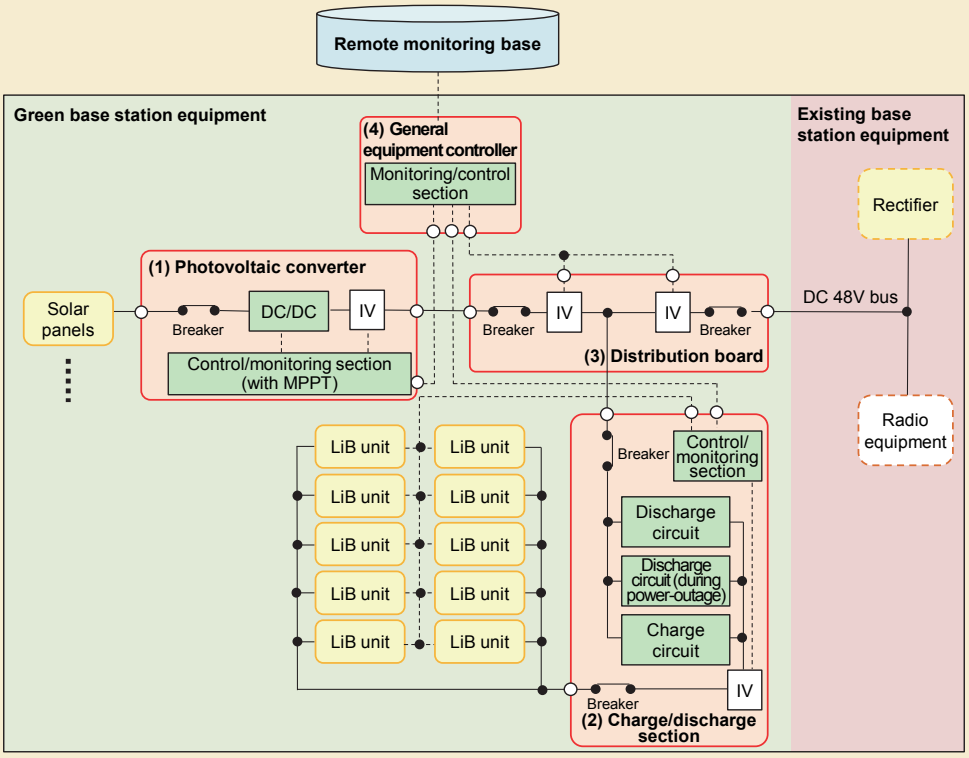
To demonstrate the effectiveness and reliability of green base stations, NTT DOCOMO has so far installed ten stations for field-testing in the Kanto-Koshinetsu region (Tokyo and Kanagawa, Gunma, Ibaraki, Yamanashi, Nagano, and Niigata prefectures). Furthermore, based on the results obtained from these stations, 44 commercial green base stations have been installed throughout

the country, which means that green base stations including the test stations are now operating in every prefecture in Japan.

A green base station improves the self-generation rate*1 by providing power shift control that efficiently stores surplus power generated in the daytime [4]. However, given that the power generated by solar panels may be insufficient or excessive depending on weather conditions, either LiB capacity must be increased or the battery's State Of Charge (SOC) must be brought sufficiently low beforehand to ensure charging without losing use of any surplus power generated. Yet, from the viewpoint of a disaster countermeasure, it is desirable to keep SOC as high as possible to ensure the provision of a power supply in the event of an emergency such as a power outage. However, increasing LiB capacity is constrained by costs and installation space, so keeping SOC low is necessary in practice, but this prevents an environmental contribution and a disaster countermeasure from being achieved together. To solve this issue, NTT DOCOMO has developed a technology called "weathercast-linked control." This article describes the configuration of a green base station, gives an overview of this technology, and reports on the results of field trials conducted at a test station.

2. Configuration of Green Base Station

A typical configuration of a green base station is shown in **Figure 1**. In addition to a rectifier*2, which serves as a power-supply facility in conventional radio base stations, a green base station introduces LiB in place of lead-acid batteries and connects the solar panels to a DC 48V bus*3 (hereinafter referred to as "DC bus") in parallel so that generated power can be used at times of a power outage. By incorporating



IV: Current/voltage measurement location. I: Symbol for current, V: Symbol for voltage.

Figure 1 Typical configuration of a green base station

*1 **Self-generation rate:** Percentage of consumed power covered by photovoltaic generation.

*2 **Rectifier:** A device for converting AC commercial power to DC power for radio equipment.

*3 **DC 48V bus:** Power line between the rectifier and radio equipment. Radio base stations use a DC 48V voltage.

a power-supply control section for high-efficiency control of these three types of power supplies, a green base station achieves effective use of natural energy and deals with peak power demand.

The power-supply control section consists of a (1) photovoltaic converter, (2) charge/discharge section, (3) distribution board, and (4) general equipment controller as described below.

(1) The photovoltaic converter incorporates a Maximum Power Point Tracking (MPPT)*4 function to generate power at maximum efficiency [5] [6]. It also sets the output voltage after Direct Current-to-Direct Current (DC/DC) conversion higher than the rectifier and LiB voltage to give priority to photovoltaic power in the supply of power to the radio equipment.

(2) The charge/discharge section incorporates charge and discharge circuits and performs a specific amount of charging and discharging based on instructions received from the control/monitoring section. The discharge circuit (during power-outage) includes a circuit for detecting voltage drops in the DC bus so that power can be supplied from the batteries without interruption at the time of a power outage.

(3) The distribution board provides a connection to the DC bus and detaches individual equipment from the DC bus at the time of an emergency.

(4) The general equipment controller performs unified management of generated power measured at IV current/voltage measurement locations and of various types of measurement data such as battery-unit information. It includes a monitoring function for passing information to a remote base and a function for responding to remotely received control instructions.

3. Overview of Weathercast-linked Control

An overview of weathercast-linked control is shown in **Figure 2**. This technology features two operation modes: sunlight control mode during normal times and autonomous/regeneration mode during power outages.

- In sunlight control mode, the process makes changes as needed to battery SOC based on information in weather reports so as to make maximum use of

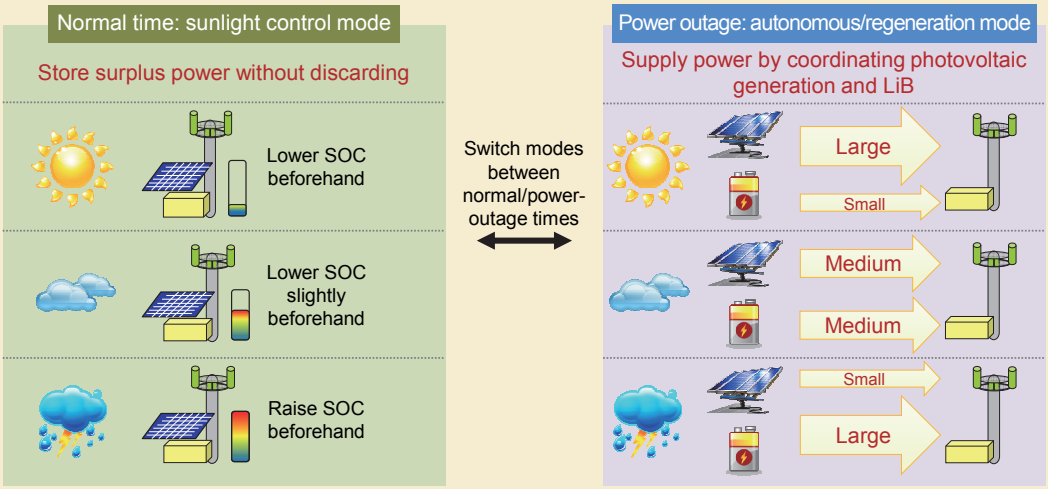


Figure 2 Overview of weathercast-linked control

*4 **MPPT:** A control technique for maximizing power generation by controlling the voltage of solar panels.

generated power. That is to say, it makes use of the fact that power generation by solar panels is highly dependent on the weather. For example, given a forecast of clear weather, SOC could be lowered beforehand at nighttime by discharging the batteries and supplying power to the radio equipment. This would enable charging in the daytime without losing use of any surplus power.

- In autonomous/regeneration mode, LiB supplies power to the radio equipment to make up for the insufficient portion of power from photovoltaic generation. Moreover, whenever surplus power happens to be generated by photovoltaic generation, the process switches to a charging operation without interrupting the above supply of power. This makes for maximum use of generated power the same as during normal times even during power outages.

These two modes make it possible to achieve both an environmental contribution and a disaster countermeasure. Specifically, if a power outage should occur after lowering SOC given a forecast of clear weather, a high backup effect can still be expected and the drop in LiB backup ability can be supplemented by making maximum use of the power generated by the solar panels the same as in normal times. Furthermore, if a power outage should occur after raising SOC given a forecast of rain, photovoltaic generation cannot be expected but the high LiB backup ability at this time can be used to compensate for the drop in solar-panel backup ability.

3.1 Sunlight Control Mode

1) Overview

Weather-dependent operation in sunlight control mode is shown in **Figure 3**. In this mode, the amount of surplus power generated by the solar panels is estimated based on weather forecasts. Then, by discharging that amount by time T when surplus power generation begins, solar-panel generated power that was conventionally discarded can be effectively used even for batteries that cannot hold sufficient capacity other than that needed for backup during a power outage. Moreover, at the time of an emergency such as a power outage, a solar-panel backup effect achieved by switching to autonomous/regeneration

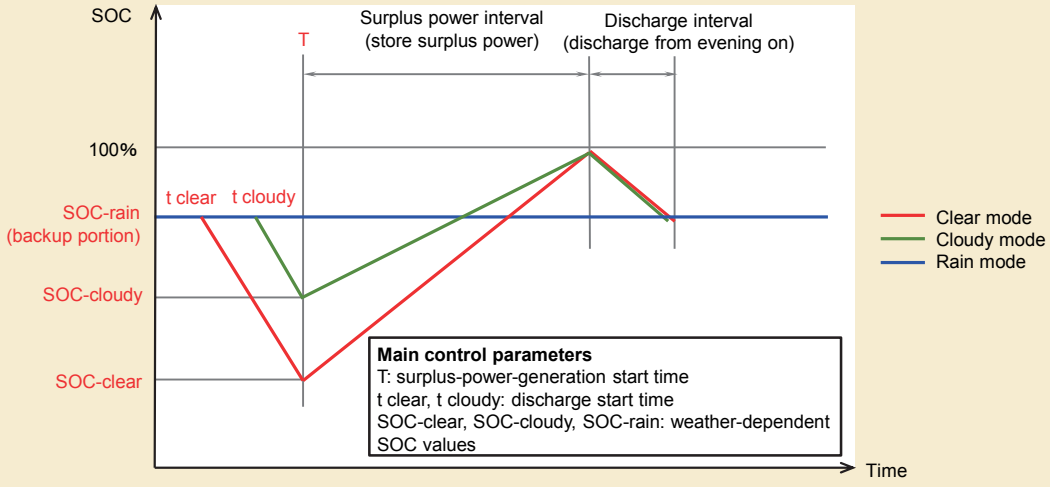


Figure 3 Operation in sunlight control mode

mode enables communication services to be continued for a specific amount of time.

2) Control Technique in Sunlight Control Mode

The following describes the control technique in sunlight control mode.

This technique begins by determining the weather-dependent SOC value (SOC-clear, SOC-cloudy, SOC-rain) and the discharge start time (t clear, t cloudy) based on information in the previous day's weather report. If rain is forecast, surplus power from the solar panels cannot be expected, so there is no need to discharge battery power beforehand. In this case, SOC-rain is determined to be the amount of backup needed for ensuring communications during a power outage. SOC-clear, meanwhile, is determined to be a value such that the maximum amount of surplus power generated in the last month is available for battery charging. Finally, SOC-cloudy is determined from the average value of SOC-clear and SOC-rain. Surplus-power-generation start time T is determined as the earliest time from among the actual start times for the last month, and discharge start time t is determined using full-charge battery capacity and the power load such that weather-dependent SOC is reached at surplus-power-generation start time T. Weather information for the base station is obtained before time t, and surplus power generated by the solar panels is used for battery charging after surplus-power-generation start time T. The evening-discharge start time is set from some time in the evening such as sunset when no surplus power is being generated, and on reaching this time, battery discharging begins so that SOC drops to SOC-rain (backup portion), at which time LiB enters a standby state. However, if the weather forecast includes the possibility of a natural disaster, LiB will be put into a fully charged state and operation will continue as usual until the disaster forecast is lifted. Additionally, if a disaster actually occurs and a state of power outage is entered, the radio equipment will be supplied with power from LiB, and when the power-outage state ends, LiB will again be charged up to the SOC-rain level.

3.2 Autonomous/regeneration Mode

1) Overview

Operation in autonomous/regeneration mode is shown in **Figure 4**. In the past, the LiB SOC would decrease monotonously at the time of a power outage due to a natural disaster or other

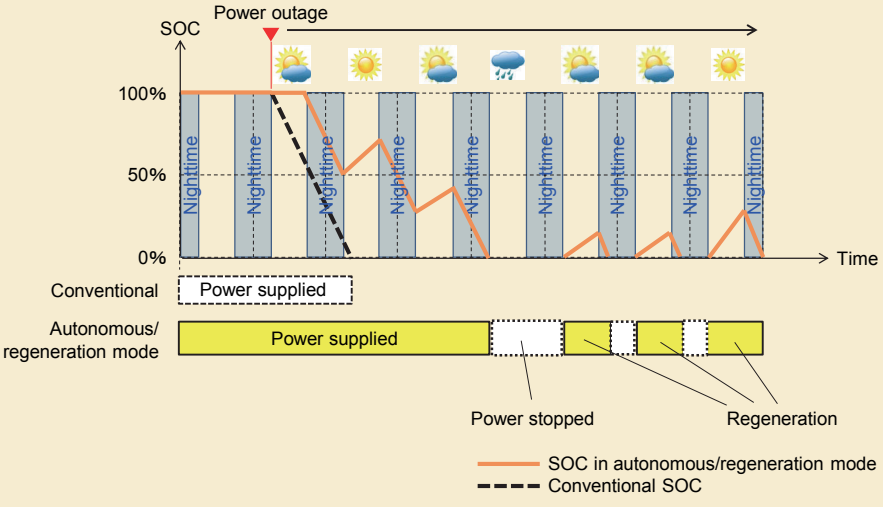


Figure 4 Operation in autonomous/regeneration mode

event, given that radio-equipment load was about constant. As a result, battery power would gradually diminish leading to a cut in service. In autonomous/regeneration mode, however, the power generated by the solar panels can be supplied to the radio equipment during a clear day and surplus power can be used to charge LiB thereby extending service time. Regeneration operation can be performed in the same way even after LiB depletes.

2) Control Technique in Autonomous/regeneration Mode

The following describes the control technique in autonomous/regeneration mode.

When using surplus power from photovoltaic generation to charge LiB at the time of a power outage, it is necessary to prevent a loss of power feeding to the radio equipment by using more power than the amount of surplus power for charging. In normal times, when charging batteries with more power than the amount of surplus power, an amount of power equivalent to this excess portion (the shortfall in power fed to the radio equipment) will be provided from the rectifier. However, in autonomous/regeneration mode during a power outage, power feeding from the rectifier cannot be expected. A technique is therefore needed to estimate the amount of surplus power so that charging using more power than that amount will not take place in autonomous/regeneration mode.

Estimating surplus power requires information on generated power and consumed power. While there is a technique for calculating generated power based on measurements of solar irradiance [7], irradiance can, in actuality, vary locally due to partial shading on solar panels and other effects, which makes reliable estimations difficult. In addition, power consumption in radio equipment can fluctuate slightly according to traffic conditions, which adds to the difficulty of making reliable estimates of surplus power.

For the above reasons, we apply a local search algorithm called the hill climbing method*5 as a technique for estimating surplus power in autonomous/regeneration mode. Specifically, when it is seen that rectifier output is zero and surplus power is being generated, this method increases the amount of charging in very small increments. Then, when it is observed that surplus power is not being generated, the method decreases the amount of charging in very small decrements. Repeating this process in such small steps enables surplus power to be estimated and only surplus power to be used for charging. Furthermore, while sudden drops in power generation can occur due to momentary shading by passing clouds on a day that is otherwise clear, the charge/discharge section can deal with this by switching to a discharge state from a charge state so that power feeding to the radio equipment is uninterrupted. The charging amount is then reset to a minimum value in preparation for the next round of charging by surplus power.

4. Field Trials

We installed a green base station providing 1.4 kW-rated photovoltaic generation at a radio base station in Gunma prefecture as a field test station. The power consumption of the radio equipment at this test station was approximately 0.5 kW. During the test, various types of measurement data regarding the power-supply system and any system alarms such as abnormal battery temperature were sent to a remote monitoring base over a communications circuit. The green base station was found to operate well with no system alarms generated during the testing period.

4.1 Test Results for Sunlight Control Mode

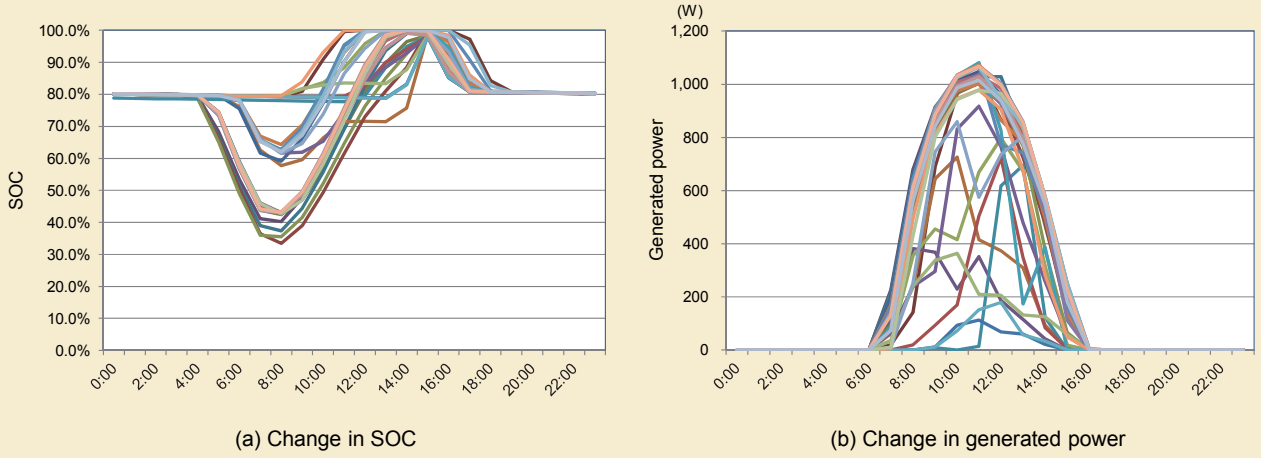
Test results for sunlight control mode are shown in **Figure 5**. For this demonstration, we conducted a long-term test from December 4, 2015 to January 3, 2016 using LiB with a capacity of 3.75 kWh. The graphs in the figure overlay the results for this period. Examining the change in SOC in Fig. 5(a), it can be seen from the manner of falling SOC by early-morning discharging that control could be performed according to the three prescribed control patterns (clear, cloudy, rain). Additionally, examining the change in generated power in Fig. 5(b),

*5 **Hill climbing method:** A search algorithm for finding the optimum of a function.

it can be seen that the amount of generated power was generally greater than the power consumption of the radio equipment during the day so that surplus power could be effectively used. These test results show that an improvement in the self-generation rate of approximately 11% can be expected compared with conventional power shift control under the same conditions.

4.2 Test Results for Autonomous/Regeneration Mode

Test results for autonomous/regeneration mode are shown in **Figure 6**. In this test, we used LiB with a capacity



(a) Change in SOC (b) Change in generated power

Figure 5 Test results for sunlight control mode

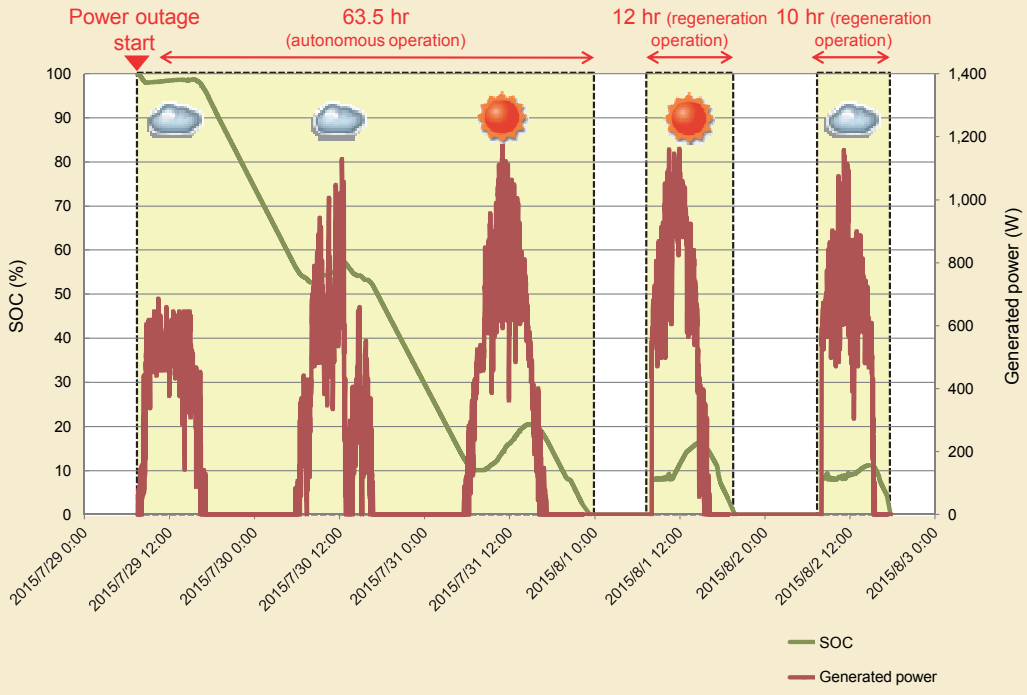


Figure 6 Test results for autonomous/regeneration mode

of 13.5 kWh and reproduced power-outage conditions by turning off the power supply to the rectifier in the field test station. The graphs in the figure show photovoltaic power generation (brown plot) and LiB SOC (green plot) after the beginning of a power outage. It can be seen that LiB charging took place even during a power outage, which means that power generation that included surplus power greater than the amount of power consumed by the radio equipment was performed. These results show that the autonomous operation period was approximately 2.4 times that when using LiB of the same capacity as an emergency power supply. It can also be seen from these results that the radio base station could operate for a certain period of time through power supplied from LiB even after sunset depending on the amount of power stored during the day.

5. Conclusion

This article described weathercast-linked control that achieves both an environmental contribution and a disaster countermeasure in green base stations. The use of this control technology can be expected to improve the self-generation rate, reduce the amount of commercial power used, and extend service time during a power outage. Going forward, we plan to enhance the service-time extension effect during power outages without diminishing the self-generation rate by further dividing control/operation modes to raise the level of LiB backup time. We also plan to promote the introduction of this technology in parallel with the expansion of green base stations.

References

[1] I. Suzuki, T. Shizuki and K. Nishiyama: "High power and long life lithium-ion battery for backup power sources," Proc. of INTELEC '03, pp.317-322, Yokohama, Oct. 2003.

[2] T. Matsushima: "Characteristics of Lithium-Ion Cells with Large-Capacity and DC-Power-Supply Systems for Telecommunications Applications," Transactions of the Institute of Electronics, Information and Communication Engineers (B), Vol.J91-B, No.12, pp.1725-1734, Dec. 2008 (in Japanese).

[3] K. Takeno: "Deployment of Green Base Station," Journal of the Institute of Electrical Installation Engineers of Japan, Vol.34, No.2, pp.110-113, Sep. 2014 (in Japanese).

[4] M. Nakamura, Y. Otomo and K. Takeno: "Development of power shift control for Green-BTS," IEICE General Conference, B-9-12, Mar. 2015 (in Japanese).

[5] G.J. Yu, Y.S. Jung, J.Y. Choi and G.S. Kim: "A novel two-mode MPPT control algorithm based on comparative study of existing algorithms," Proc. of Solar Energy, Vol.76, No.4, pp.455-463, 2004.

[6] T. Tafticht, K. Agbossou, M.L. Doumbia and A. Cheriti: "An improved maximum power point tracking method for photovoltaic systems," Proc. of Renewable Energy, Vol.33, No.7, pp.1508-1516, Apr. 2008.

[7] T. Suzuki, Y. Goto, T. Terazono, S. Wakao and T. Oozeki: "Forecast Method of Solar Irradiance with Just-In-Time Modeling," IEEJ Transactions on Power and Energy B, Vol.131, No.11, pp.912-919, Nov. 2011 (in Japanese).

Energy-saving Location Estimation | Location-based Content Delivery | Customer Behavior Analysis

Technology Reports

Behavior Sensing Technology for O2O Marketing —Energy-saving Location Estimation, Flexible Information Delivery Control, Customer Behavior Visualization and Analysis—

O2O marketing services are gaining attention. These services promote sales and encourage people to visit actual shops based on customer location information acquired through smartphones. Issues with the O2O marketing service include how to combine high-accuracy user location estimation using various smartphone sensors with reduced power consumption, provide information delivery opportunity controls and perform marketing through analysis of customer behaviors.

*NTT DOCOMO has developed behavior sensing technology that estimates position using abstracted map information and provides flexible information delivery controls to respond to customer behaviors etc., and visualization for customer behavior analysis. Including some of these technologies in Air Stamp®*1 enables timely and energy-saving location-based information delivery and detailed customer behavior analysis.*

Research Laboratories
Hiroyuki Sano
Teppei Konishi
Satoshi Kawasaki
Masakatsu Tsukamoto
Smart-life Planning Department
Masaru Maebatake†

1. Introduction

Using the sensors built into smartphones enables provision of services linked to indoor/outdoor customer location information. These services can encourage shop visitation by delivering electronic coupons for shops in the customer's current vicinity to the customer's smartphone and deliver in-store product recommendations. Predicted to expand, the size of the market for these "Online to Offline" (O2O) marketing services*2 could reach approximately JPY 236 billion by 2020 [1].

O2O marketing services aim to encourage customer shop visitation, migration and purchase by providing interaction with store information, product information and coupons etc. With this technology, we believe it will be possible to more effectively achieve the objectives described above by providing content matched to customer preferences and needs, and matched to customer locations and their presence in areas in a timely manner.

† Currently Corporate Strategy & Planning Department

*1 **Air Stamp®:** A registered trademark of NTT DOCOMO, INC.

*2 **O2O marketing services:** Information services which aim to influence customer purchasing behaviors in actual stores by making use of interaction with information on the Internet.

As basic technology for providing O2O marketing services, NTT DOCOMO has developed a behavior sensing system that integrates developments for (1) energy-saving location estimation, (2) flexible information delivery controls, and (3) visual analysis of customer behaviors. This system has been included in the "Air Stamp" [2] location information platform provided by NTT DOCOMO. This article presents O2O marketing usage cases, and describes an overview of the behavior sensing system as well as examples of deploying it into Air Stamp.

2. Usage Cases

In providing O2O marketing using behavior sensing system, we assume the three usage cases of encouraging customer store visitation, migration and sales promotions in commercial facilities such as shopping malls, department stores and train station buildings in which there are many shops as well as shops such as supermarkets, home improvement centers and convenience stores. We assume that privately owned smartphones (hereinafter referred to as "terminals") will be the main destination for information delivery, although tablets on shopping carts in supermarkets or home improvement centers could also be targeted.

- Store visitation promotion: The system delivers information when customers are in the vicinity of commercial facilities or shops using outdoor location information. The system also provides customers with a welcome message upon entering a store, and delivers a survey upon leaving.
- Migration promotion: To get customers to visit more shops in commercial facilities or in-store points of sale, the system delivers information on shops and in-store points of sale they have not yet visited or shop and point of sale recommendations. Like a stamp rally, the system delivers a coupon when the customer has visited all designated shops.
- Sales promotions: Using indoor location data, the system delivers information matched to the point of sale or product display shelving. This involves replaying videos about products directly in front of product display shelving to attract customer attention, and displaying coupons available for use if the customer is at a point of sale for a certain amount of time, and is unsure about making a purchase.

Commercial facilities and shops can view visualizations of customer shop visitation behaviors and in-store movements to uncover issues, and go through the "Plan Do Check Act" (PDCA) cycle*3 to consider measures (plan), enact them (do), measure results (check), and improve content and timing of delivered information (act) to maximize effectiveness of measures etc.

3. Behavior Sensing System

3.1 System Configuration

To realize a system to handle the aforementioned usage cases, functions for (1) energy-saving location estimation, (2) flexible information delivery control and (3) customer visitation behavior and in-store movement visual analysis are required. We developed a behavior sensing system to provide and integrate these three functions. **Figure 1** describes an overview of the system.

(1) Energy-saving location estimation functions (Fig. 1 (1)): To estimate terminal location, the system generally acquires values from various sensors in terminals such as GPS, Wi-Fi®*4, Bluetooth® Low Energy (BLE)*5 and makes various computations on the values acquired to convert them to location information. Conventionally, because a system has to do location estimation regularly and often by sensor scanning to continually determine customer indoor/outdoor location or presence, and determine detailed customer positioning at the level of in-store product display shelving and the time that customers are present in an area, there is an issue with increased load on the customer's terminal battery. Therefore, since

*3 **PDCA cycle:** A method of ensuring smooth running of business. The PDCA cycle entails repeatedly and continually running through the four steps of (1) Plan (planning), (2) Do (performing), (3) Check (measuring results) and (4) Act (making improvements).

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

*5 **BLE:** An extension of the Bluetooth short-range radio communications specifications, added in Bluetooth version 4.0, and characterized by its low-power communications. Bluetooth is a registered trademark of Bluetooth SIG Inc. in the United States.

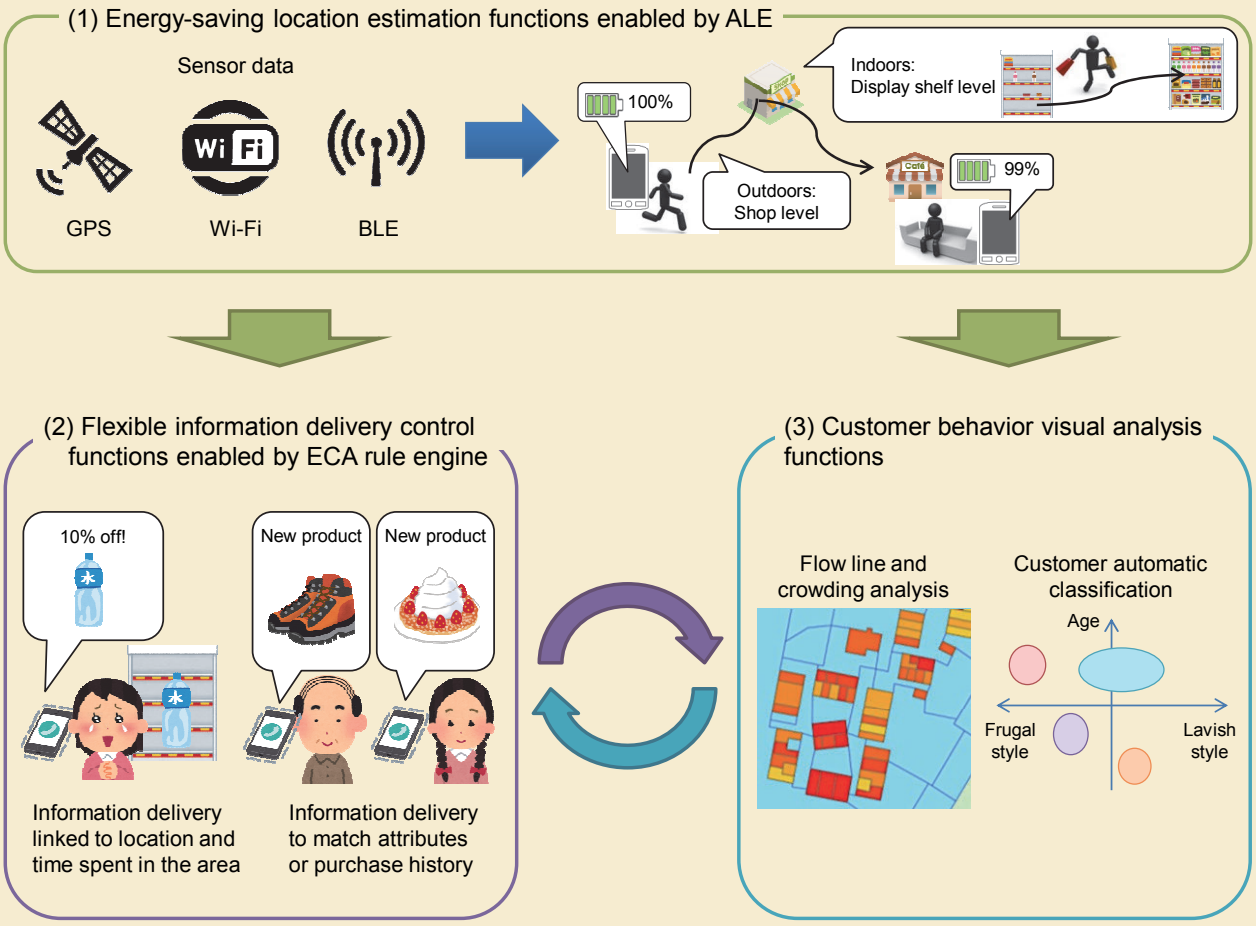


Figure 1 Overview of behavior sensing system

customers may spend several hours shopping in a commercial facility, the amount of power consumed by the terminal during that time must be reduced. However in general, it is possible to lengthen the interval that data is acquired from sensors to lower power consumption but location estimation accuracy will be lower as a result due to a trade off between power consumption and estimation accuracy [3]. To solve this issue, we developed a location estimation engine, called "Abstract Location Engine (ALE)." It runs on the terminal, and combines energy saving with location estimation accuracy at the product display shelving level in-store, and at the shop level outdoors. ALE achieves this using an abstracted map with a node and link configuration (described later) to compensate for location estimation errors caused by more infrequent acquisition of values from sensors, and performs path interpolation.

(2) Flexible information delivery control functions (Fig. 1 (2)): For flexible information delivery, the system requires functions to analyze customer attributes such as age and gender available from store membership card registrations, and purchasing history acquired from store Point-Of-Sale (POS) information to deliver information to each customer tailored to their needs and preferences, and linked with their location or presence in an area. Further, the system has to be able to handle demands from commercial facilities and shops to quickly perform new measures or change the content or timing of delivered information while measuring the effectiveness of measures. These demands can be met with Event Condition Action (ECA) rules. ECA rules consist of event, condition and action statements to control information delivery discussed in the aforementioned usage cases. An ECA rule engine runs on the terminal to execute the ECA rules, and compares the stated delivery rules with terminal states such as location data, time data and application operation logs to perform operations such as displaying content or notifying a server when specified conditions are met. Because ECA rules set in the server are immediately reflected in the application, it is possible to quickly handle new measures or change measures without updating the terminal application.

(3) Customer behavior visual analysis functions (Fig. 1 (3)): To study measures and measure effectiveness, the system requires functions to visualize customer store visitation and in-store behaviors such as flow lines and crowding. The customer behavior visual analysis functions include functions for real-time display of basic information such as numbers of customers who visited shops, numbers of repeat visitations and the time that customers were present based on specified dates, times or demographic attributes*6, as well as analysis based on automatic customer classification using individual flow lines, crowding conditions and flow patterns.

These three functions are described in details as follows.

3.2 Energy-saving Location Estimation Functions Enabled by ALE

1) ALE Overview

ALE is incorporated into applications running on the user's terminal to estimate the location of the terminal using the combined values of various sensors such as GPS, Wi-Fi and BLE.

(1) Node and link configuration

ALE corrects location estimation errors due to infrequent sensor value acquisition and noise in measured sensor values using map information abstracted with the node and link configuration, and then outputs results of location estimation in abstracted units called "nodes."

The abstracted map used by ALE displays areas for location estimation such as shops or commercial facilities as a graph structure with nodes and links. An example of an abstracted map of a commercial facility converted into nodes and links is shown in **Figure 2**.

- Setting nodes

O2O marketing service providers set locations where it is desirable to acquire information such as in front of shelving displaying sale items or at intersections in aisles as nodes in the abstracted map. The aforementioned information includes user behaviors such as passing traffic or presence in areas. There are two types of nodes, indoor nodes and outdoor nodes. User presence in outdoor nodes is determined from GPS positioning. At indoor nodes where GPS radio waves are unavailable, user presence is determined from the signal strength of short-distance radio communication devices such as Wi-Fi access points

*6 **Demographic attributes:** Information from population statistics used to characterize customers. These include attributes such as gender, age, job and number of persons in the family unit.

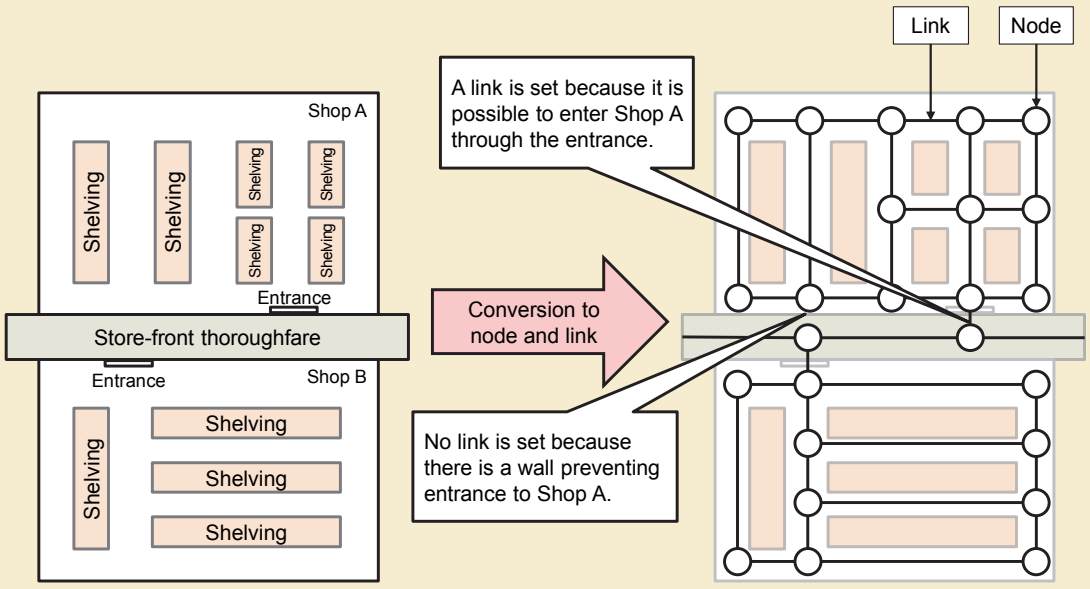


Figure 2 Example of abstracted map with node and link configuration

and BLE tags set up in the area.

• Setting links

The links in the abstracted map represent aisles, and are only set for routes between nodes through which people can physically and directly move. For example, as shown in Fig. 2, there is no direct route between Shop A and Shop B, which means the only passage is through the nodes in the thoroughfare in front of the shops. Hence, there are no direct links connecting nodes in Shop A with those in Shop B. Instead, there is a link connecting Shop A entrance with the thoroughfare in front of the shops, and a link connecting the thoroughfare in front of the shops with the Shop B entrance. There are also weights set for links. These weights represent the cost required for people to move node to node, where the larger the weight, the longer the time it takes to move. An example of a weight value could be the direct physical distance between the centers of nodes. These weights are used to correct location estimations, as described below.

(2) Location correction using node and link information

Using this node and link information, ALE suppresses the degradation of location estimation performance. The time required to transition between nodes is calculated using the link weights and average speed of movements of people. ALE compares node transitions results with positioning results to determine whether there are any abnormalities with the location estimation results and hence prevents erroneous location estimation results from being output.

(3) Node interpolation

There is a higher chance of events occurring in which user presence at nodes cannot be detected with ALE due to the longer sensor value acquisition interval to lower power consumption. Hence, when ALE detects a node transition that is not a possible direct transition on the link configuration, its node interpolation function provides a supplement by predicting the optimal path based on the link structure.

2) ALE Evaluation Experiments

As described above, while saving energy, ALE performs the above location correction and node interpolation using abstracted map information to accurately estimate location at the shelving level indoors and at the shop level outdoors. In testing performed in FY 2014, we installed ALE in a commercial facility application and performed experiments to evaluate power consumption. We confirmed that it contributes to an approximate 36% to 40% power consumption reduction, and that the location estimation performance is hardly degraded at all even when the sensor value acquisition interval is doubled.

3.3 Flexible Information Delivery Control Functions Enabled by ECA Rule Engine

1) ECA Rule Engine Overview

An overview of the ECA rule engine is shown in **Figure 3**. As a function that runs on the terminal, this interprets and executes an eXtensible Markup Language (XML)*7 file that contains the ECA rules describing the event, condition and action statements. As shown in **Table 1**, events such as changes to ALE location estimations or terminal screen display are used as opportunities to execute ECA rules.

Figure 4 shows the flow of ECA rule processing. When a change occurs, the ECA rule engine checks whether

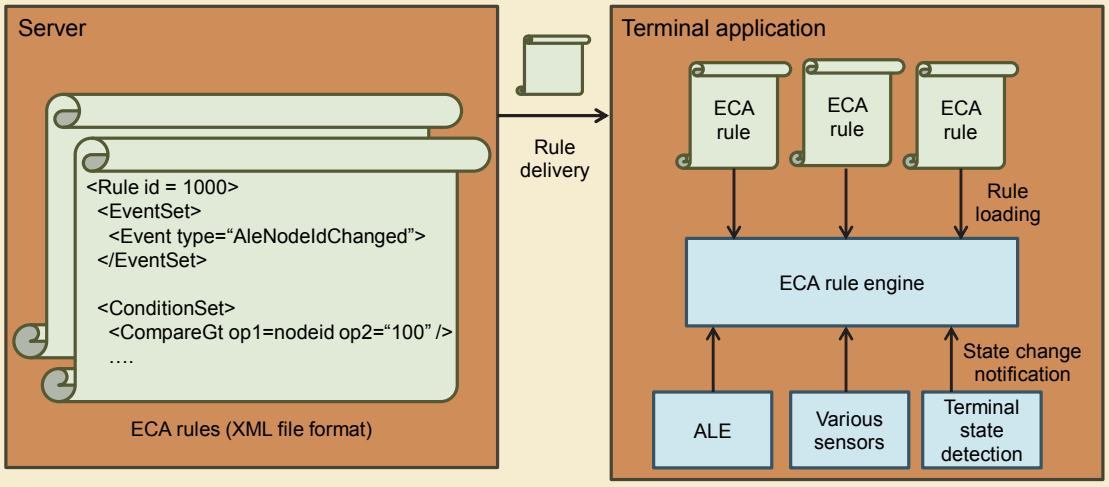


Figure 3 ECA rule engine overview

Table 1 ECA rule engine event list

Event list (partial)
Screen change notification
Screen unlock
Change to node estimated by ALE
Enter/leave registered area (region)
GPS-acquired coordinate change

*7 **XML:** A markup language for describing the meaning and structure of text and data advocated by the World Wide Web Consortium (W3C). It can be expanded, and users can specify their own unique tags.

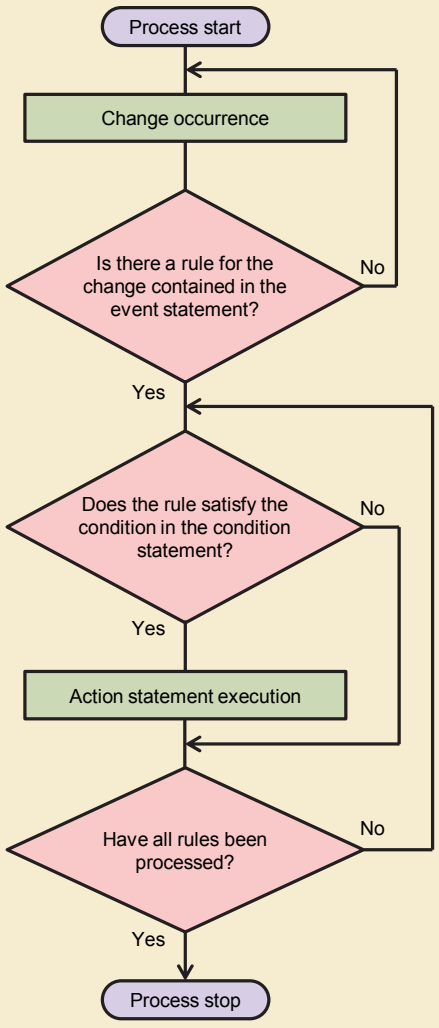


Figure 4 ECA rule process flow

the change is an event described by the event statements in the ECA rules. Events that correspond to ECA rules are determined by the conditional statements described in **Table 2**. Conditions are determined not only as current conditions by comparing with node IDs estimated by ALE or terminal status, but also compared to variables set for other ECA rule actions and so forth so that comparisons with past conditions are also possible. When conditions are satisfied, the processing described in the action statement is performed. Action statements specify identifiers for content to be displayed, and can be used to provide information to suit conditions. Apart from displaying specified content, actions such as notifying a server of shop visitation events or setting conditional variables for numbers of visitations can be specified, as shown in **Table 3**.

ECA rules are highly versatile like a kind of programming language. The ECA rule engine has a mechanism to acquire ECA rules from a server and synchronize them. Service providers can describe ECA rules for campaign measures and deliver them to terminals so that they can immediately implement measures without updating or redistributing applications. They also can perform different measures for individual users by switching ECA rules to be delivered for user IDs.

2) Usage Cases

We describe two usage cases for ECA rules.

- In the first usage case, a user is unsure about making a purchase from a product display shelf and a discount coupon is delivered to prompt the user to make a decisive purchase. This is achieved by describing an ALE estimated node change in the event statement, presence at the node for the product display shelving longer than a specified amount of time in the condition statement, and coupon content display in the action statement.
- In the second usage case, a special coupon is delivered when a user visits all specified shops, as

Table 2 ECA rule engine condition list

Condition list (partial)
A variable and specified constant are equal
A variable is lower than specified constant
Two specified variables are equal
All specified conditions are satisfied
At least one specified condition is satisfied

Table 3 ECA rule engine action list

Action list (partial)
Content display
Server notification
Enable/disable specified ECA rule
Region registration
Set value to variable

in a stamp rally. In ECA rules, when it is determined that a specified shop has been visited from ALE location estimation, the state of whether there was a visit to the shop is retained by the action statement, which writes a special variable. Then, in the condition statement, it is possible to control special coupon delivery by determining whether all specified shops have been visited.

3.4 Customer Behavior Visual Analysis Functions

These functions include functions for real-time display of basic information such as numbers of customers who visited a shop, numbers of repeat visitations and the time that customers were present based on specified dates, times and demographic attributes, as well as the following three special functions.

(1) Individual flow line visualization: Chronological analysis of flow line characteristics of individuals
(2) Crowding level visualization: Analysis of user presence in areas by each gender/age group/date and time/flow line pattern
(3) Visualization of flow line pattern characteristics: Automatic analysis of flow line patterns

The following describes usage examples of these functions.

(1) Visualization of individual flow lines enables understanding of the paths that customers follow from entering a store until leaving it, and their reaction to delivered information. For example, this makes it possible to compare individual customer paths before and after delivering information on certain point of sale specials upon store entry or around points of sale.
(2) Visualization of crowding enables display of customer presence at nodes as a heat map on the store map in which areas with many customers present are displayed in red and areas with few customers are displayed in blue. Taking into account these traffic conditions, this function enables planning, holding and measuring effectiveness of point of sale events on certain days of the week or at certain time slots, repositioning staff

to handle crowding at cash registers, and evaluation of the changing amounts of time that customers are present in areas.

(3) Visualization of the characteristics of flow line patterns entails classification of user flow lines using document classification technology, and display of them with Word clouds*8. Documents are classified according to the frequency of occurrence of words in the document. The names of products on display at each node are associated with each node in advance, a node names log with each store visit is viewed as a document, and node names in a document are viewed as words, and Latent Dirichlet Allocation (LDA)*9 is used to classify flow lines into clusters. **Figure 5** shows the automatically generated cluster 2 for flow lines of a user who frequently presents at nodes where there are side dishes, bread, fried foods and bento boxes, and displays these points of sale in the upper right of the heat map in red, indicating a high presence of the user. This means measures such as sales promotions to appeal in particular to cluster 2 customers at the points of sale in the upper right, prompting them to visit other areas that they do not visit much, or changes to store layout could be studied based on this analysis.

4. Implementation Technology into Air Stamp

We have included the behavior sensing system ALE in the "Air Stamp" location information platform provided by NTT DOCOMO. Air Stamp has been designed for shops such as supermarkets or restaurants, and commercial facilities to develop services converging the real with the Internet by linking location information so that coupons, and information about events and building directions can be delivered to users based on smartphone location estimations.

This behavior sensing technology enables well-timed and energy-saving location-based information delivery and detailed analysis of customer behaviors. As shown in **Figure 6** (a), on shelving displaying a range of different products, information delivery pin-pointed for certain products on the display is possible in front of the shelving. In addition, as shown in the train station example in Fig. 6 (b), corrections can be made for flow lines that cannot actually exist in commercial facilities to raise the accuracy of customer movement analysis. We are proceeding with studies on the release of Air Stamp supporting these

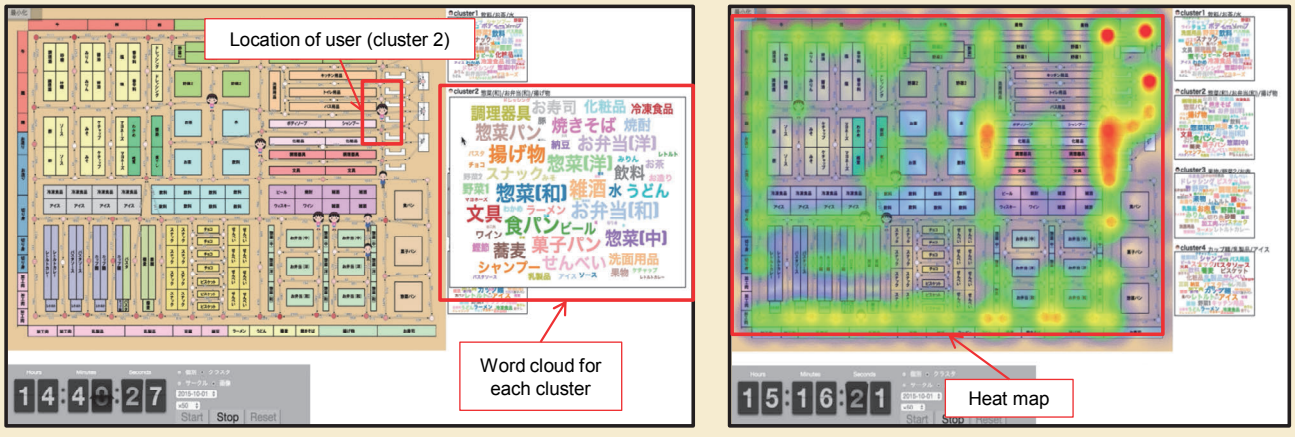
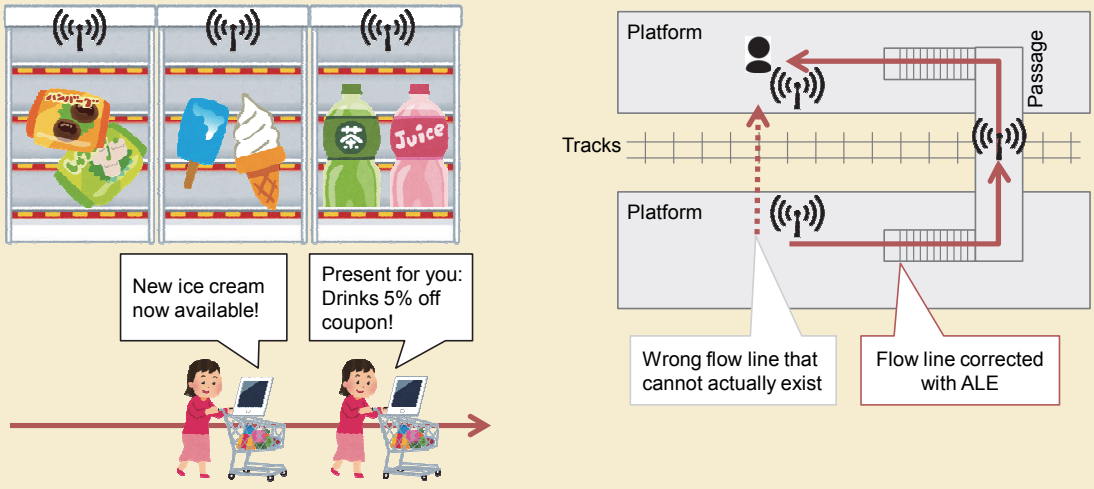


Figure 5 Customer behavior visualization and analysis functions

*8 **Word cloud:** A visualization method in which the size of words expresses the frequency of their appearance. The larger the characters, the more frequent the appearance of the word.

*9 **LDA:** A machine learning technique that classifies documentation based on the topics related to words that appear in the documentation.



(a) Location-based information delivery

(b) Customer behavior analysis

Figure 6 Location-based information delivery (a) and more accurate customer behavior analysis (b) using this behavior sensing technology

functions.

5. Conclusion

This article has described behavior sensing technologies that integrate functions required for O2O marketing services, and an overview of the system. Using this system, commercial facilities and shops will be able to quickly implement the PDCA cycle for various measures.

We plan to incorporate knowledge and data gained from the results of deploying Air Stamp services in various commercial facilities and shops into customer behavior analysis and use it to study how to make system operations more efficient.

References

[1] Seed Planning Inc.: "Current state of O2O business and future directions, 2013," 2013 (In Japanese).
[2] Air Stamp Web site. http://www.airstamp.jp
[3] H. Liu, H. Darabi, P. Banerjee and J. Liu: "Survey of Wireless Indoor Positioning Techniques and Systems," IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS - PART C: APPLICATIONS AND REVIEWS, Vol.37, No.6, pp.1067-1080, Nov. 2007.

VTC2016-Spring Best Paper Award

On May 17, 2016, General Director Takehiro Nakamura and Tetsuo Imai of 5G Laboratory, Research Laboratories at NTT DOCOMO, and Yuichi Kakishima and Haralabos Papadopoulasi from DOCOMO Innovations, Inc. were presented with the Best Paper Award at the 83rd IEEE Vehicular Technology Conference (VTC2016-Spring) held in Nanjing, China. The paper for which the award was presented was "5G 3GPP-like Channel Models for Outdoor Urban Microcellular and Macrocellular Environments" [1]. This paper is related to the 5th Generation mobile communications system (5G) evaluation channel model jointly developed with universities and external companies under the guidance of NTT DOCOMO (Aalto University, BUPT, CMCC, Ericsson, Huawei, Intel, KT Corporation, Nokia, NYU WIRELESS, Qualcomm, Samsung, University of Bristol, University of Southern California, AT&T).

The channel model models radio channel propagation characteristics required for designing mobile communications systems, and focuses on modeling modern characteristics such as channel latency, angle of departure/arrival and polarization in addition to channel loss. Models cited include the IMT Advanced Model standardized by the ITU-R as a 4th Generation (4G) model, and the 3D Channel Model standardized by 3GPP. New considerations for 5G include:

- Application of higher frequencies—Targeting frequencies from 6 GHz to 100 GHz in addition to existing frequencies
- Array antenna technology advancements—Support for Massive MIMO technologies etc.
- Diversification of system construction scenarios—scenarios including D2D (Device to Device) or V2V (Vehicle to Vehicle) for addition to existing scenarios, which were also studied as new requirements for the channel model

To efficiently develop a channel model for 5G systems, NTT DOCOMO has been driving studies with the above universities and companies over nearly two years since 2014. The results of these studies are summarized in the "5G Channel Model for bands up to 100 GHz" white paper [2]. This channel model is characterized by its high affinity with the existing 3GPP channel model and modeling based on lots of actual and simulation data from work done by partner universities and companies.

The paper for which this award was presented was appraised at VTC2016-Spring for its particular focus describing a model for outdoor urban macrocell and microcell environments from the aforementioned white paper. A model for indoor environments was also reported at IEEE ICC2016 [3]. Please refer to the white paper for details about these channel models.

References

[1] K. Haneda, L. Tian, Y. Zheng, H. Asplund, J. Li, Y. Wang, D. Steer, C. Li, T. Balercia, S. Lee, Y. Kim, A. Ghosh, T. Thomas, T. Nakamura, Y. Kakishima, T. Imai, H. Papadopoulas, T. S. Rappaport, G. R. MacCartney Jr., M. K. Samimi, S. Sun, O. Koymen, S. Hur, J. Park, C. Zhang, E. Mellios, A. F. Molisch, S. S. Ghassamzadeh and A. Ghosh: "5G 3GPP-like Channel Models for Outdoor Urban Microcellular and Macrocellular Environments," VTC 2016-Spring, May 2016.
http://arxiv.org/abs/1602.07533

[2] 3rd Workshop on Mobile Communications in Higher Frequency Bands (MCHFB): "White paper on "5G Channel Model for bands up to 100 GHz."
http://www.5gworkshops.com/5GCM.html

[3] K. Haneda, L. Tian, H. Asplund, J. Li, Y. Wang, D. Steer, C. Li, T. Balercia, S. Lee, Y. Kim, A. Ghosh, T. Thomas, T. Nakamura, Y. Kakishima, T. Imai, H. Papadopoulas, T. S. Rappaport, G. R. MacCartney Jr., M. K. Samimi, S. Sun, O. Koymen, S. Hur, J. Park, J. Zhang, E. Mellios, A. F. Molisch, S. S. Ghassamzadeh and A. Ghosh: "Indoor 5G 3GPP-like Channel Models for Office and Shopping Mall Environments," IEEE ICC2016-Workshop, pp.694-699, May 2016.

Certificate of Appreciation Received from NGMN Alliance

In March 2015, the Next Generation Mobile Networks (NGMN) Alliance published a white paper on the vision and requirements for the 5th Generation mobile communication networks (5G) [1]. On June 21, 2016, NGMN awarded a Certificate of Appreciation to Dr. Mikio Iwamura of DOCOMO Communications Laboratories Europe GmbH, representing NTT DOCOMO, in appreciation of NTT DOCOMO's proactive contributions to the creation of this white paper.

NGMN was established in 2006 by the world's leading telecommunications operators with a mission to define operator visions and requirements for next-generation mobile communications systems. As a founding member, NTT DOCOMO has been contributing to NGMN activities to continue development of mobile communications systems as a platform to support various activities in society and foster new communications culture.

Currently, 28 operators from around the globe are participating in NGMN, working toward realization of 5G, together with 36 leading global vendors as contributors and 25 university and academic institutions as advisers. After the launch of the NGMN 5G Initiative in March 2014, 5G was discussed from various angles over the course of a year, and the outcomes were summarized in the white paper. Notably, DOCOMO Communications Laboratories Europe led the group discussing 5G technologies and system architecture, and made significant contributions to establish 5G technical concepts and design principles, and build consensus in the industry. These contributions were highly appraised by NGMN which led to the awarding of this certificate, along with the other leading key members including Deutsche Telekom, Bell Mobility, AT&T, Telecom Italia, Telefonica, China Mobile Communications, Orange, Vodafone Group and British Telecommunications.

Embracing the outlook for technologies and the changing business environments for operators, the NGMN 5G white paper highlights the following points as its 5G vision for building the telecommunications industry ecosystem*1 for 2020 and beyond.

- Support increasing data demands in society and enable seamless user experience.
- Enable secure and trusted ICT environments that are indispensable to improve productivity in society.

- Support the paradigm shift to IoT with the power of ICT.
- Support efficient realization of diverse services using software and cloud computing technologies.
- Foster value creation through partnerships that combine capabilities.

To support these directions, NGMN advocates expansion of capabilities such as data rates, capacity, latency and the number of devices that can be connected, as well as network architecture that enables agile and flexible tailoring of capabilities and network functions to suit various services. This will require integration of a new radio access technology to further improve spectral efficiency and unlock new frequency bands such as millimeter waves, and flexible cloud technologies that exploit virtualization and software.

To integrate various technical concepts into a sensible system in a timely manner, intimate knowledge of technologies and business insights are needed. To facilitate a global 5G ecosystem, the NGMN 5G Initiative successfully synthesized the wisdom of various operators and crystallized the vision into its white paper.

The NGMN 5G white paper has been recognized globally as a key foundation in building the 5G ecosystem, and global standardization and commercial developments are underway toward realization of its vision. Also, from June 2015 to June 2016, NGMN further reviewed 5G requirements from the business [2] and vertical industry [3] perspectives, and studied detailed radio access requirements [4] and so forth published in various white papers. NTT DOCOMO also took initiatives as leader, facilitator and key contributor in these activities, to take the realization of 5G a step further.

References

[1] NGMN: "5G White Paper." https://www.ngmn.org/fileadmin/ngmn/content/downloads/Technical/2015/NGMN_5G_White_Paper_V1_0.pdf
[2] NGMN: "5G Prospects - Key Capabilities to Unlock Digital Opportunities." Jun. 2016. https://www.ngmn.org/uploads/media/160701_NGMN_BPG_Capabilities_Whitepaper_v1_1.pdf
[3] NGMN: "Perspectives on Vertical Industries and Implications for 5G." Jun. 2016. https://www.ngmn.org/uploads/media/160610_NGMN_Perspectives_on_Vertical_Industries_and_Implications_for_5G_v1_0.pdf
[4] NGMN: "Recommendations for NGMN KPIs and Requirements for 5G." Jun. 2016. https://www.ngmn.org/uploads/media/160603_Annex_-_NGMN_Liaison_to_3GPP_RAN__72_v1_0.pdf

*1 **Ecosystem:** A symbiotic mechanism in which multiple companies partner across specialties to make use of each other's technologies and assets, and in which consumers and society are included in business activities to create flows from research and development through to sales, advertising and consumption, and enable coexistence and co-prosperity.

NTT DOCOMO
Technical Journal Vol.18 No.3

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