

DOCOMO DriveNet Info

*NTT DOCOMO has developed a colorful new in-vehicle-support application called “DOCOMO DriveNet Info TM*1,” which gives information generated in the cloud that is convenient while driving, such as traffic congestion and other local information, simply by speaking to the smartphone. This article gives an overview of this application as well as the speech interaction function and ITS cloud system provided.*

M2M Business Department

Ryohei Kurita

Mei Hasegawa

Service & Solution Development Department[†]

Hiroshi Fujimoto

Communication Device Development Department

Kazuaki Takahashi

1. Introduction

NTT DOCOMO Inc. (DOCOMO) and Pioneer Corp. (Pioneer) have jointly developed a new in-vehicle-support application called “DOCOMO DriveNet Info,” which gives information generated in the cloud that is convenient while driving, such as traffic congestion and other local information, simply by speaking to the smartphone. The cloud service used by this application began operating as a DOCOMO service on December 18, 2013 (the application can be used by downloading from GooglePlayTM*2 or the dmenu).

This service is DOCOMO’s Intelligent Transport System (ITS)*3 (Figure 1), which uses Pioneer’s next-

generation vehicle-oriented cloud infrastructure called “Mobile Telematics Center,” to provide real-time-updated traffic and other information convenient for driving, to smartphones. By combining with the speech understanding and speech synthesis technologies used in “Shabette Concier,” various functions such as checking traffic conditions, local information, schedules or the latest news, placing a phone call, sending and receiving SMS*4, or playing some music can be used by simply talking to the smartphone.

In this article, we give an overview of the DOCOMO DriveNet Info application as well as the speech interaction function and ITS cloud system that support it.

2. The DOCOMO DriveNet Info Appli

The cloud service used by DOCOMO DriveNet Info assumes it will be used from within a vehicle.

Because of this, the application will be used mainly through voice commands, and the GUI is quite simple. Also, to make operation more comfortable, the application links with a Smartphone Holder 01*5, and the speech-input UI can be invoked without touching the screen through the Near-Field Communication (NFC)*6 automatic launch feature of the Holder 01, and accessory switches on the steering wheel.

The application is able to access three types of information through

©2014 NTT DOCOMO, INC.

Copies of articles may be reproduced only for personal, noncommercial use, provided that the name NTT DOCOMO Technical Journal, the name(s) of the author(s), the title and date of the article appear in the copies.

[†] Currently Service Innovation Department

*1 **DOCOMO DriveInfoTM**: A trademark or registered trademark of NTT DOCOMO, INC.

*2 **GooglePlayTM**: A service from Google for delivering applications, video, music and books to Android terminals. Google PlayTM is a trademark or registered trademark of Google, Inc. U.S.A.

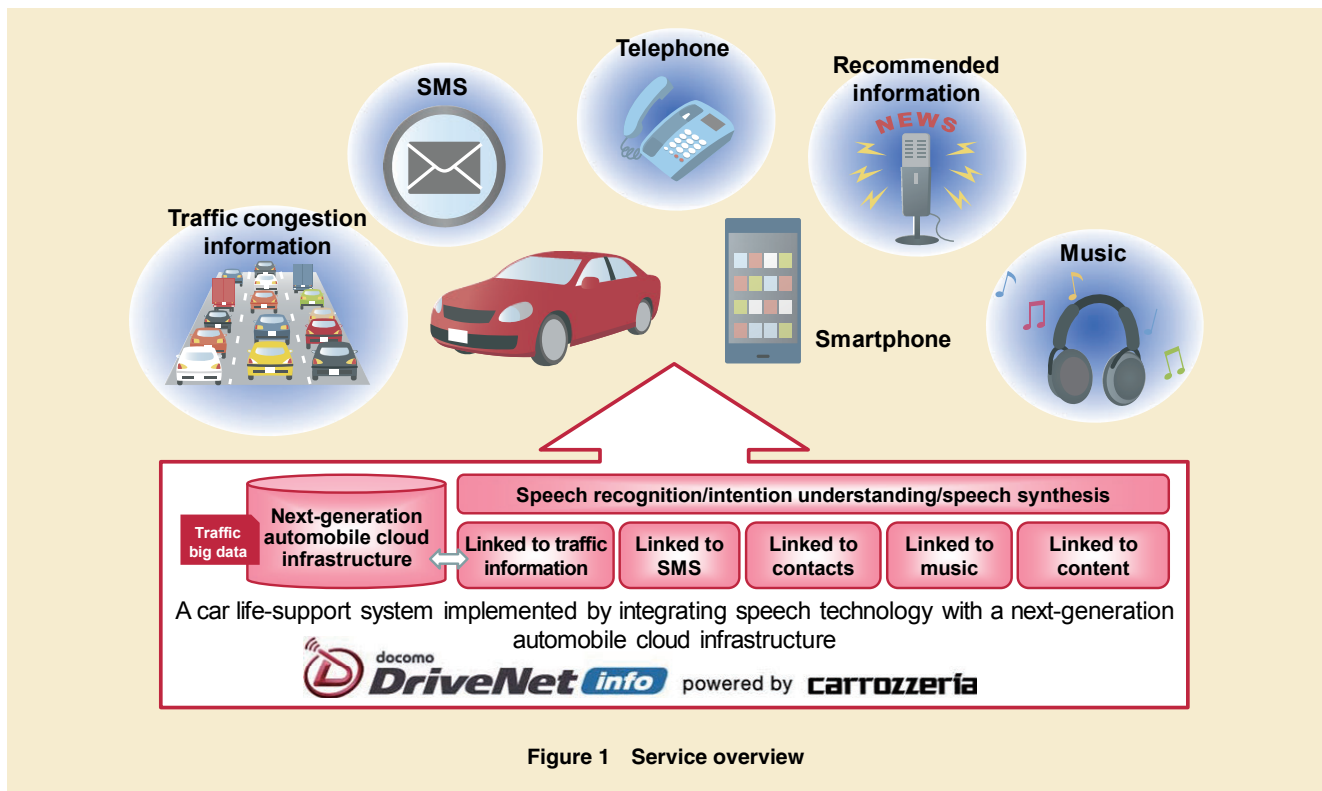


Figure 1 Service overview

speech operations.

- (1) Local terminal information
- (2) Static information on the network (contacts list, facility information, etc.)
- (3) Dynamic information on the network (real-time traffic information)

1) Local Terminal Information

Functions such as playing music on a playlist or reading out SMS messages are implemented using local terminal information.

2) Static Information on the Network

Functions such as placing a phone call or sending an SMS or e-mail message from the contacts list, or identifying the originator of a phone call, SMS or e-mail message, are done using static

information on the network. The application also provides a character that reads out a variety of information useful during a morning commute in a morning-news program format (**Figure 2**). This includes:

- Gasoline prices
- Schedules
- Weather information
- News and current affairs
- Horoscope
- Restaurants
- Traffic information

Search for nearby facilities, or free keyword searches can also be done, destinations can be selected from the search results, and the estimated time of

arrival at the destination can be checked. Route information can also be obtained by linking with the DOCOMO DriveNet Navi™^{*7} appli. The search function uses the speech interaction function, and destinations can be found using a conversational process (**Figure 3**).

3) Dynamic Information on the Network

Dynamic information on the network includes traffic congestion information. Location information collected by Pioneer car navigation products, DOCOMO DriveNet Info, and DOCOMO DriveNet Navi (iOS Version) is analyzed on the cloud and distributed as traffic congestion data from the mobile telematics center. Congestion information can also be

*3 **ITS**: An overall name for transportation systems using communications technology to improve vehicle management, traffic flow and other issues.

*4 **SMS**: A service for sending and receiving short, text-based messages, mainly between mobile terminals. It can also be used for sending and receiving control signals to mobile terminals.

*5 **Smartphone Holder 01**: A holder for securing a smartphone inside a vehicle. Supports NFC. Includes a power cable for the smartphone and remote control for speech recognition on the steering wheel.

*6 **NFC**: A contactless communication technology like FeliCa®, and the related standard. FeliCa® is a registered trademark of Sony Corp.

*7 **DOCOMO DriveNet Navi** ™: A car navigation application provided since April, 2011. DOCOMO DriveNet Navi ™ is a trademark and registered trademark of NTT DOCOMO Inc.

viewed on a map screen customized for car navigation.

When information that is more current is needed, there is also a traffic information query function. The set of users driving ahead of the user on the same road can be queried regarding whether there is traffic congestion (**Figure 4**). When the server receives a request from a user for such confirmation, it performs a survey of users driving ahead of the user on the same roadway, and these users can reply to the survey with a simple “yes” or “no”. The information from these users is processed statistically on the server and distributed together with user-posted congestion information. In addition to user requests, the server also determines where new congestion might be occurring from other information sent to the server (driving speed, etc.), and initiates its own surveys with users in the area regarding whether there is congestion (**Figure 5**).

To perform these functions, it is important that vehicle position information be real time. To achieve this, not only is the information sent periodically to the server at all times, but information is sent dynamically, according to the driving conditions, to ensure the server has accurate information.

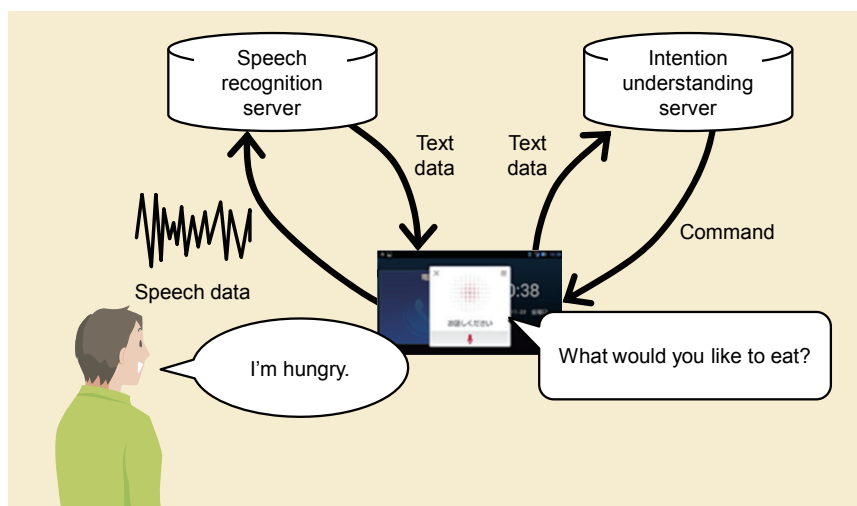
3. Speech Interaction Function

DOCOMO DriveNet Info has a



※This is provided only in Japanese at present.

Figure 2 Example of a speech notification from DOCOMO DriveNet Info



※This is provided only in Japanese at present.

Figure 3 Conversational destination search

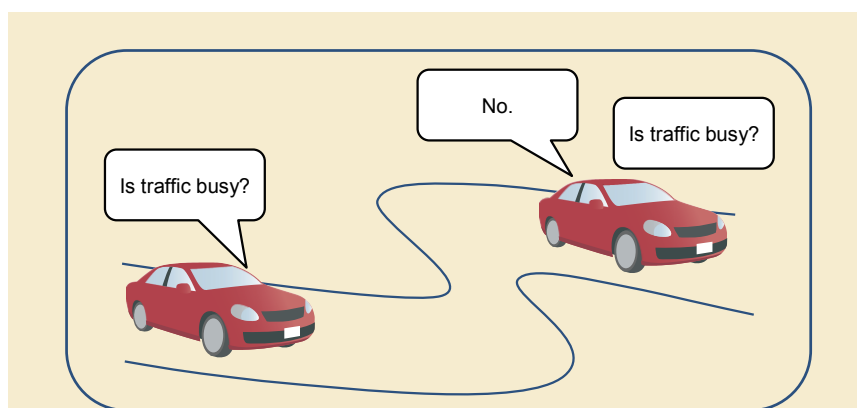


Figure 4 Checking traffic conditions

speech interaction function that enables various processes to be performed through speech interaction between user and appli, as shown in **Table 1**. For example, the process of placing a phone call is shown in **Figure 6**. The system organization for the speech interaction function is shown in **Figure 7**. The new components developed by DOCOMO on the speech interaction server include the intention understanding server and the ITS cloud system. The speech interaction function is implemented with the following process.

- (1) Speech recognition of user utterances
- (2) Understanding the intentions of user utterances
 - (2)*Get further information through interaction
- (3) Get the response for the application

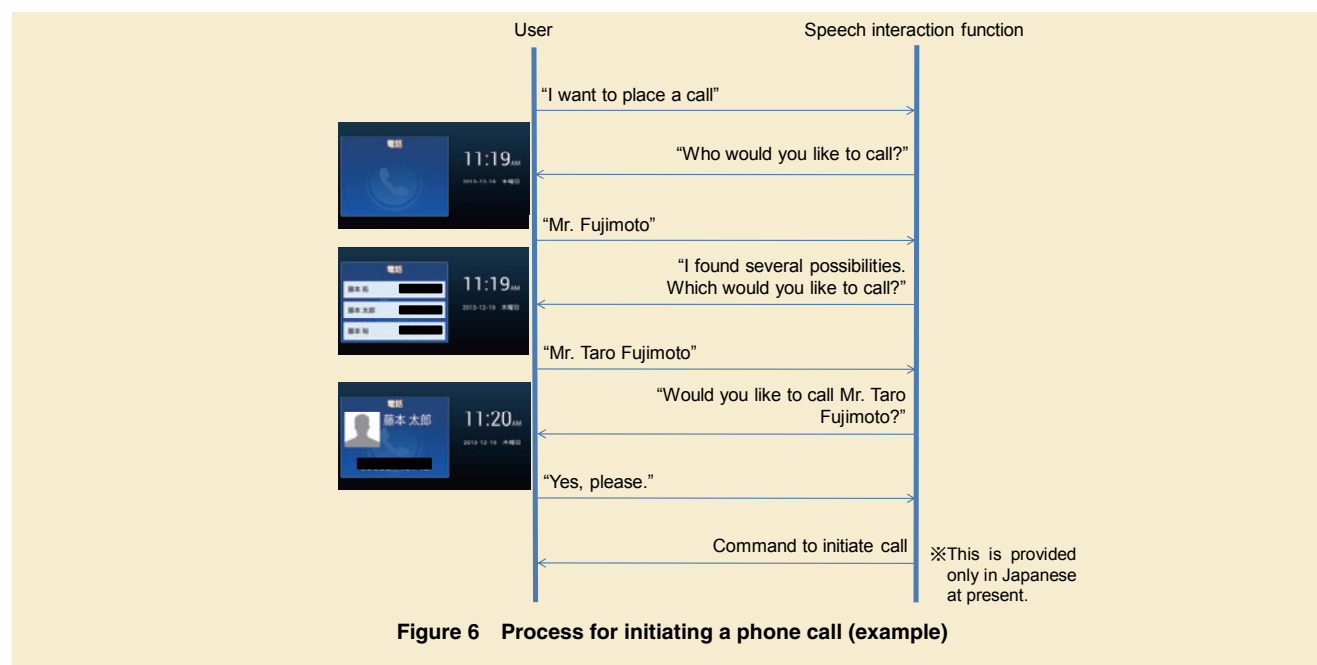


※This is provided only in Japanese at present.

Figure 5 Checking traffic conditions from the server

Table 1 Processing that can be implemented with conversation

Type	Details	Process
Map processing	Search for place names or facilities	Search for a destination
	Search for restaurants	Search for a restaurant
Contact list processing	Place a phone call SMS	Phone
	Send an SMS	Send an SMS
Chat processing	Respond to chat.	Chat
Operation processing	Zoom in/out on map	Zoom in/out on the map
	Re-play the news	Re-play the news



※This is provided only in Japanese at present.

Figure 6 Process for initiating a phone call (example)

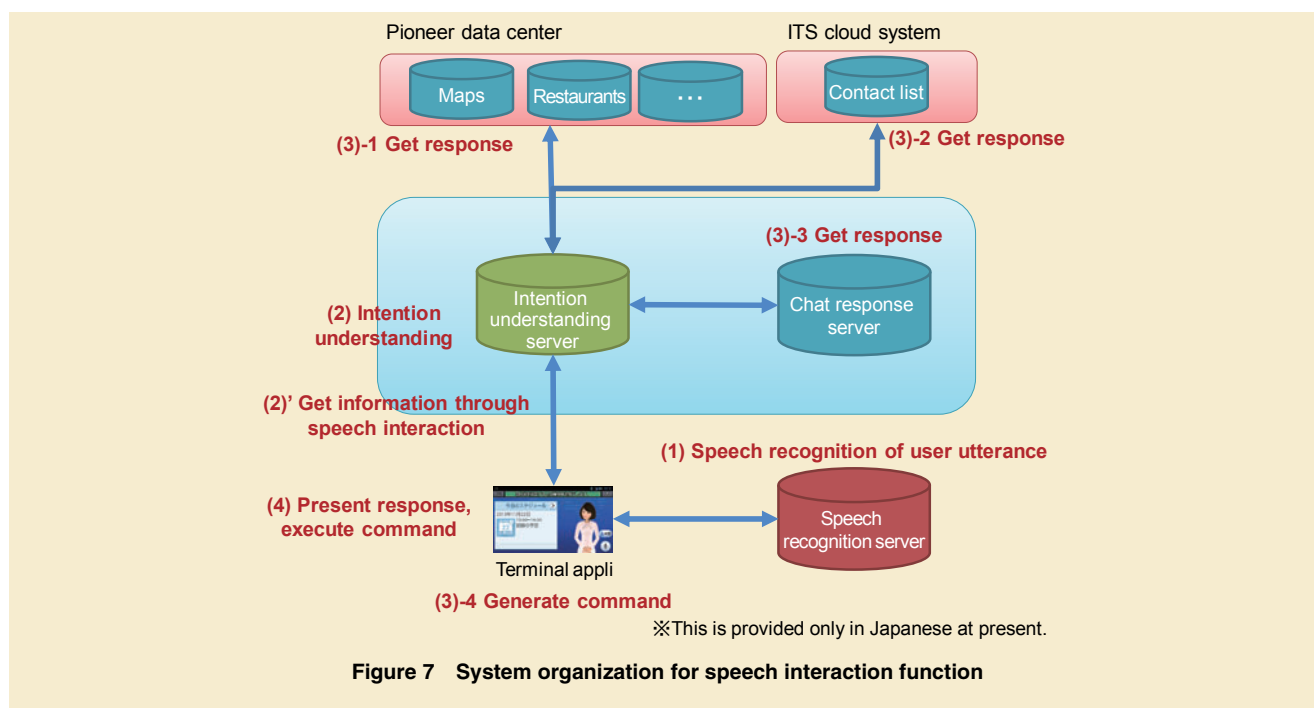


Figure 7 System organization for speech interaction function

or generate a command
(4) Present the response or execute the command

Speech recognition of the user utterance is done from the application through the DOCOMO speech recognition UI application^{*8} by the speech recognition server (Fig. 7 (1)). The speech recognized by the server is converted to text and then sent to the intention understanding server through the appli.

The intention understanding server determines what sort of task was intended by the user's utterance (Fig. 7(2)). This is essentially a text-classification problem that classifies the input text. Specifically, a learning model based on many sample utterances for each

possible task is used to quickly determine the type of task the input text was intended to initiate, as shown in Figure 8 [1].

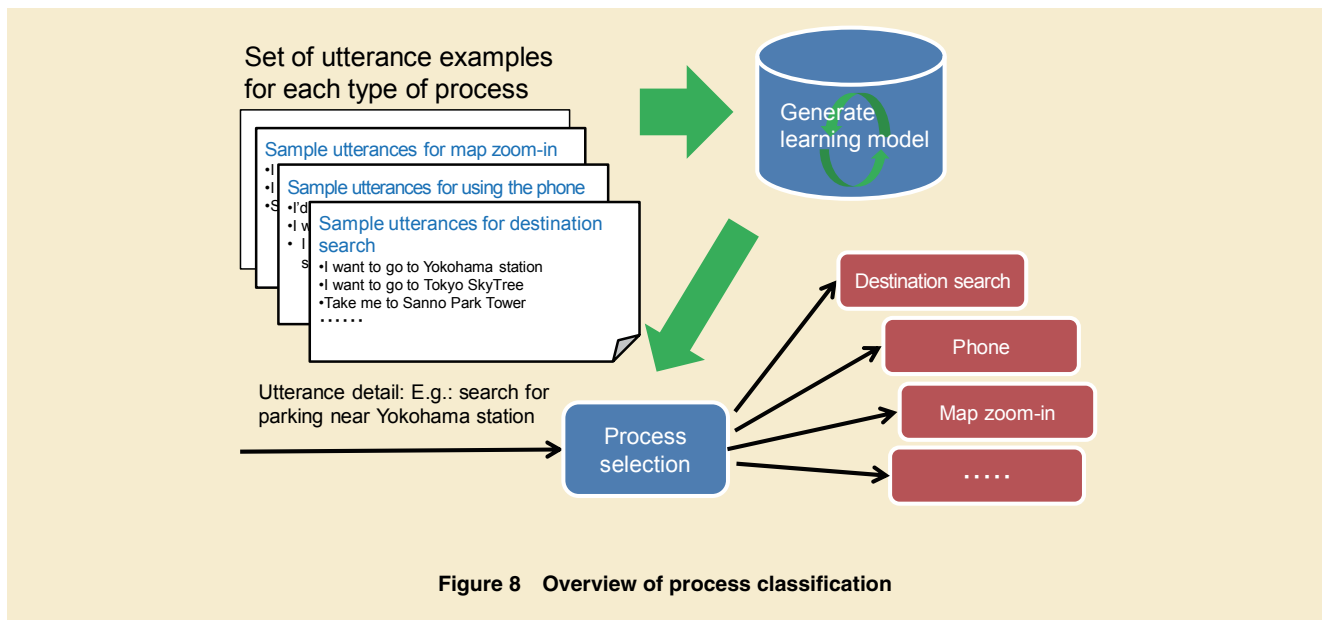
Here, there can be cases when there is not yet enough information to perform the selected task. For example, to process a call, the person to call must be known, so if the person was not included in the user's utterance, the task cannot be performed. Also, since DOCOMO DriveNet Info assumes the operations are being done inside a vehicle, it cannot respond in a way that requires the user to look or use his/her hands, such as displaying the address book on the smartphone screen.

Accordingly, the speech interaction function of DOCOMO DriveNet Info fills in the information needed for the

task through speech interaction with the user (Fig. 7 (2)'). For example, with the example shown in Fig. 6, the destination is filled in. If all of the information required to perform the task cannot be extracted from the user's utterance, the speech interaction function asks the user again, and gets the needed information from the user's response.

The tasks selected by the intention understanding function are classified into four main categories as shown in Table 1. If the result of classification is map processing, such as a destination or restaurant search, answers needed to respond, such as location information, are retrieved from the Pioneer data center (Fig. 7 (3)-1). If the task is to use the contacts list to initiate a phone call or SMS message, answers such as the

^{*8} **UI application:** An application that runs on the mobile device, providing an interface for the user. On Android terminals in particular, it is a Java application. Oracle and Java are registered trademarks of Oracle Corporation, its subsidiaries and affiliates in the United States and other countries.



destination address are retrieved from the contact list server of the ITS cloud system (Fig. 7(3)-2). For chat processing, responses are retrieved from the chat response server (Fig. 7(3)-3)[2]. Finally, for operations such as zooming the map, a simple command is generated, to be executed by the application (Fig. 7(3)-4).

The responses and commands retrieved as above are sent, through the intention understanding server, to the application to perform each process (Fig. 7(4)).

4. ITS Cloud System

An overview of the ITS cloud system is shown in **Figure 9**.

The ITS cloud system for DOCOMO DriveNet Info has the following functions.

1) Probe Data Management Function

Probe data management is a secure

data management function for handling sensitive big data, including location information.

The probe data sent by the smartphone application is classified as personal information. Because of this, it must be encrypted and stored in a way that individuals cannot be identified before use as big data, and encryption consumes a certain amount of resources. On the other hand, large amounts of information are processed in real time, and rapid responses are required. The ITS cloud system must handle these conflicting requirements. Specifically, the device ID of data received from the application is hashed^{*9}, and secure, high-speed processing is implemented by using a real-time in-memory DB^{*10}. All data is also encrypted when storing the in-memory DB in a regular DB.

This function is also designed with

the ability to scale out^{*11}, in order to handle future increases in traffic.

2) Contact List Management Function

The contact list management function manages data for using the speech interaction function to place calls and send e-mail messages.

To phone or e-mail with users in the contacts list, the ITS cloud system maintains a contacts list, performs searches through the intention understanding server, and determines the person to be phoned or e-mailed using speech interaction. Searching incurs a processing load, depending on the number of items in the contacts list and the search keywords, so it must be performed by the ITS cloud system. To do this, users must synchronize their contact list with the server on the network.

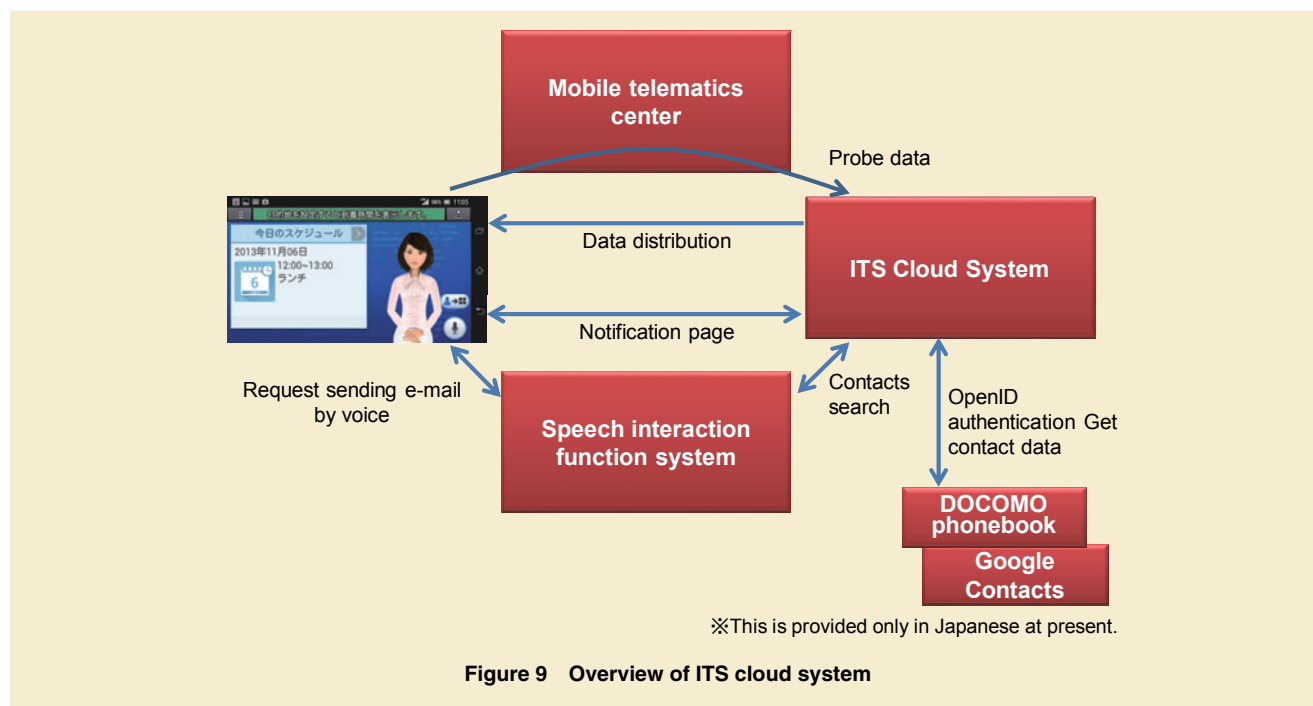
Currently, the DOCOMO phonebook^{*12}

^{*9} **Hashing:** Use of a hash function to compute a hash value from the original data. Note that it is extremely difficult to compute the original data from the hash value.

^{*10} **In-memory DB:** A database that expands and maintains the data in memory to accelerate the response of database access.

^{*11} **Scale-out:** Adding and assigning new resources to increase processing capacity when there is insufficient processing capacity on the network due to increasing service requests.

^{*12} **DOCOMO Phonebook:** The NTT DOCOMO Phonebook appli. It is backed up on a cloud server and can be used from multiple devices.



from DOCOMO and Google Contacts^{*13} from Google can be used. The ITS cloud system performs ID authentication (OpenID authentication) in order to use the phonebook, and the phonebook data is stored in a secure DB.

3) Information Distribution Function

The information distribution function distributes information to the appli. Information distributed can be selected based on factors such as device, area, or specific roads.

For example, when requesting congestion information, the user queries vehicles ahead of him/her whether there is congestion. The mobile telematics center and ITS cloud system do this by searching for vehicles to be queried based on location, road ID, and travel

direction, and the ITS cloud system distributes the query information to the appli.

4) Notification Page

When the DOCOMO DriveNet Info application is launched, a notifications screen is displayed, showing content stored in the cloud infrastructure and provided to the application by the Web server function.

5. Conclusion

In this article, we have described the DOCOMO DriveNet Info system. It combines speech intention understanding and speech synthesis technologies to provide information regarding traffic congestion, the surrounding area, schedules or the latest news, and perform tasks

such as placing or receiving phone calls, sending or receiving SMS, or playing music, simply by speaking to a smartphone. We plan to improve this user experience further using probe information and other big data in the future.

REFERENCES

- [1] K. Onishi et al.: "Casual Conversation Technology Achieving Natural Dialog with Computers," NTT DOCOMO Technical Journal, Vol. 15, No. 4, pp. 16-21, Apr. 2014.
- [2] H. Fujimoto, T. Hara and S. Nishio: "Development of Car Navigation System Operated by Naturally Speaking," IEICE transactions on information and systems, Vol. J96-D, No. 11, pp. 2815-2824, Nov. 2013 (in Japanese).

^{*13} **Google Contacts:** A contact list managed in a Google account. The data is stored on a server and can be used on multiple devices through the Google account.