

# Development of Voice Based Services In an Open Telematics Environment

Chul-Su Kim, Jongik Kim, Daesub Yoon and Hyun-Suk Kim  
Telematics·USN Division  
Electronics and Telecommunications Research Institute  
161 Gajeong-dong, Yuseong-gu, Daejeon, 305-350 KOREA

*Abstract:* - Telematics is a converging technology to provide users in a car with various services using communication devices and computers such as navigation service, point-of-interest service, email service etc. It is very difficult and dangerous for drivers to input data into a small device and read something while they are on the road. In this paper we suggest a system architecture for developing of voice based services in open telematics environment to keep eyes on the road, hands on the wheel. We developed voice based telematics services based on telematics gateway and framework for open telematics system.

*Key-Words:* -Telematics, voice, speech, CTI, vehicle

## 1 Introduction

Vehicle telematics is a converging technology to provide a driver with various services (e.g. navigation, point-of-interest, email, etc.) using communication-enabled computing devices. To make use of telematics services, however, a driver must interact with the telematics terminal device, which can cause serious safety problems. To solve the safety problem, it is indispensable that drivers should be able to keep their attention on driving while they interact with the telematics terminal device. For this purpose, a driver can input a command using speech and telematics terminal device can output the service information using voice channel. In this paper, we investigate how to implement a telematics system that enables a vehicle user to interact with a telematics terminal device using speech and sound.

We can categorize a telematics system using speech and sound into two groups. The first is that a terminal device handles speech recognition and output sound creation. When a user issues a command with speech, the terminal device convert the speech into the command text using a speech recognition engine on it. When some information should be transferred to the user, the terminal device converts the information into sound using a text-to-speech engine. The second technique is that a server-side service providing system handles speech recognition and output sound creation. We can use the well-known CTI (computer-telephony integration) technique. A vehicle user issues a command using speech and the speech

sound is transferred to the CTI center using a telephone line. Then the CTI center transfers the command sound to either a speech recognition engine or a human operator. After some information for the user request is created, either a server-side text-to-speech engine or a human operator read the information and transfer the output sound to the terminal device using a telephone line.

Because of low computing power of a terminal device, the first approach can fail speech recognition in many cases, and the sound created in a terminal device has low quality. Thus, the speech recognition and sound creation in the server-side is preferred to the terminal-based approach. However, the server-side approach requires the complex CTI infrastructure which is very expensive. Also, a terminal device should support both telephone and data channel communication in the server-side approach.

In this study, we developed a new telematics system that provides telematics service with speech and sound in the server-side. Our system uses only data channel communication and does not require complex CTI infrastructure, while it provides high quality telematics service with speech and sound. This paper is organized as follow. In Section 2, we describe some related work. In Section 3, we present development of voice based services using an open telematics system solution. In Section 4, we show two telematics services developed in this study. Finally, we conclude the paper in Section 5

## 2. Related Work

### 2.1 Open Telematics System

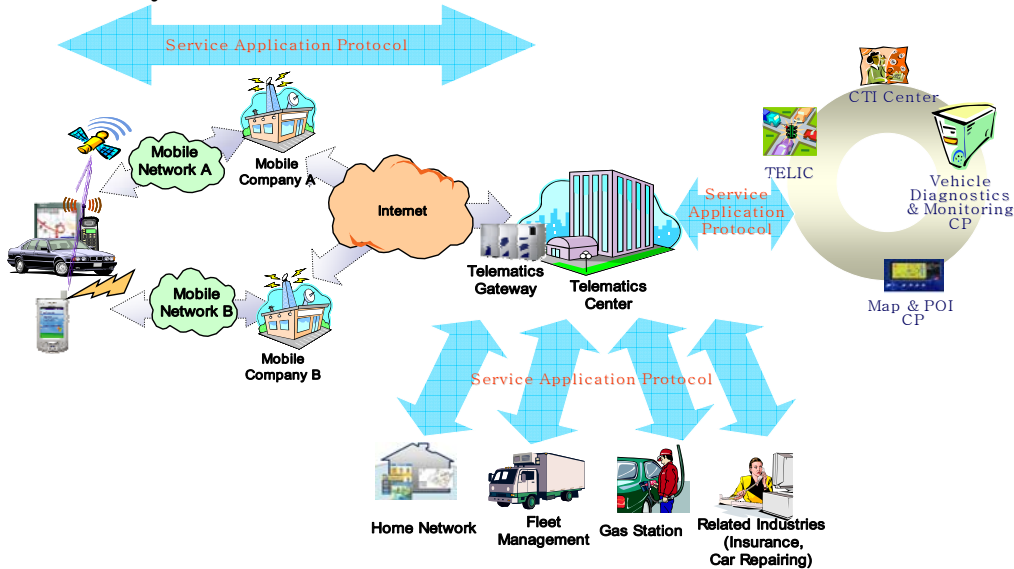


Fig.1. Open Telematics System

Most telematics systems using a mobile network for data communication use WAP (Wireless Application Protocol) gateway in the network to mediate data transmission between terminal devices in wireless network and service servers in wired network. In this case, services are dependent on the network. To resolve this dependency problem, the telematics system such as the open telematics system [1] proposed to locate telematics gateway in a telematics service center as shown in Fig. 1.

Telematics service providers generate and maintain contents for their services. But it is a big burden to a small company and it becomes an obstacle for the company to begin new business in telematics fields. Therefore open telematics system proposed the technology which enables them to make a telematics service with various contents distributed in the Internet. We adopted the open telematics system for the base of our system, which enables service providers to use voice service center as one of the external content providers. Therefore they can provide voice based services without setting up voice processing systems in their companies.

### 2.2 Telematics & voice related technologies

Telematics has become more popular and widespread with the rapid evolution of information technology and automotive industry. Most research has been conducted in technical aspects [1] such as

developing network environments and driver performance research toward safety issues [2 and 3] such as how usage of cellular telephones in telematics environments affect human attention and performance.

In driving environments, safety is one of the most important concerns so telematics services should not disturb drivers' attention while drivers are driving. According to Wickens' multiple resource theory [4], hands-free telematics operation would be predicted to be safer than hand-held operation. This result shows the possibility of the use of speech based interfaces in telematics environment to increase safety by reducing drivers' distraction.

Speech-based interfaces can be one of possible candidate to mitigate drivers' distraction, and the speech recognition system should be reliable and accurate. With respect to the performance of speech recognition, McMillan and Eggleston [5] found out speech recognition is slowly accepted due to technology limitations and high user expectations. To resolve these problems, we decided to analyze speech data at the server instead of analyzing the data in the telematics terminal devices. It demands high computing cost to have high performance speech recognition engines in the telematics terminal devices. The cost can be reduced if we can use server-based speech recognition engine via network.

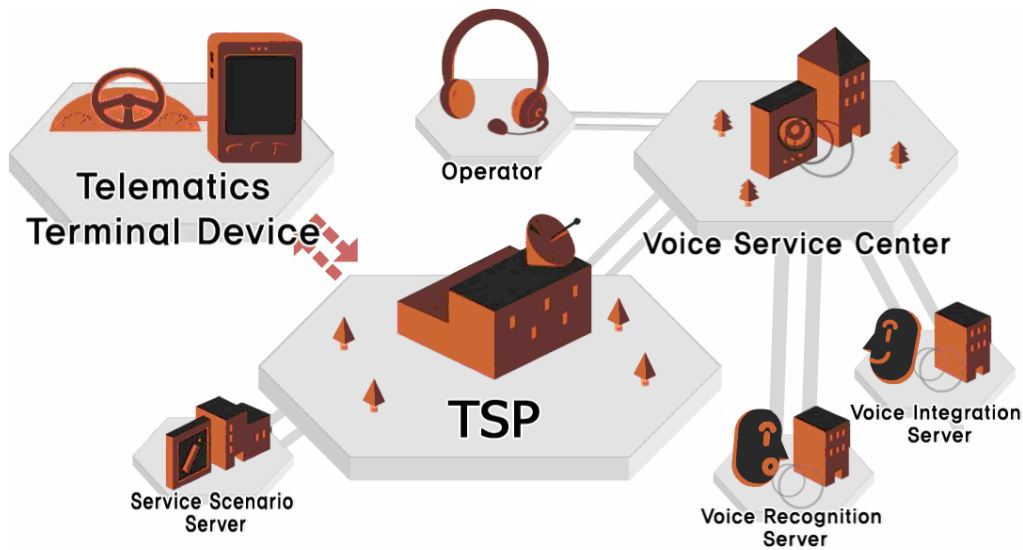


Fig.2. Voice based telematics service system

Recently, Wargnier [6] found out speech recognition acceptance rates drop off from 95% to 82% when the vehicle is at 70 mph. To implement 100% speech recognition acceptance rates, we added the functionality of operator into our system. When the server-based speech recognition engine fails to recognize drivers' voice, the recorded sound is automatically transferred to operator center. After the operator listen to the recorded voice, she responses to drivers' voice correctly. In the drivers' point of view, the system provides 100% speech recognition acceptance rates.

### 3. Voice Based Telematics Service System Development

Our voice based telematics service system has three distinguishable features compared to other telematics systems. First, most voice based service systems have their own voice processing engine, but our system does not have voice processing engine in the telematics terminal device. The telematics terminal device sends the speech data to the voice center through TSP and the voice service center sends the speech data to voice recognition server. The voice recognition server analyze actual speech data. The telematics terminal device can provide high performance of speech recognition without burden of computing cost via voice recognition server. The framework of this system is shown in Fig.2. Second, the voice service center has an operator who listens to the voice and response correctly when the voice recognition server fails to recognize the speech data.

While driving a car, it is very dangerous to ask a driver to input command repeatedly due to voice recognition fail. Our system will provide 100% speech recognition acceptance rates from drivers' point of view. Third, we used VoiceXML technology to develop the voice based service system. Service scenarios and a service processor are separated in the different system, so the voice based service processor does not need to be modified much when service scenarios are changed or new services are needed.

Telematics services typically use mobile networks between vehicles and telematics service providers because they are provided to a special environment that is a moving car. We chose a CDMA network which is most popular wireless networking technique in Korea to provide reliable services to the high speed moving car.

We developed a voice based email service for telematics which is necessary while driver's eyes are on road. Using the email service, a driver gives a command with his voice to check email list, and listen to the list and the content of email. We also developed a traffic and POI information providing service which is one of the most useful services for drivers.

#### 3.1 Framework based telematics system

The developers can write applications without knowing about details to integrate the related servers distributed in network using APIs which the framework supports. The framework supports various connection components which can connect telematics server applications into external related server applications.

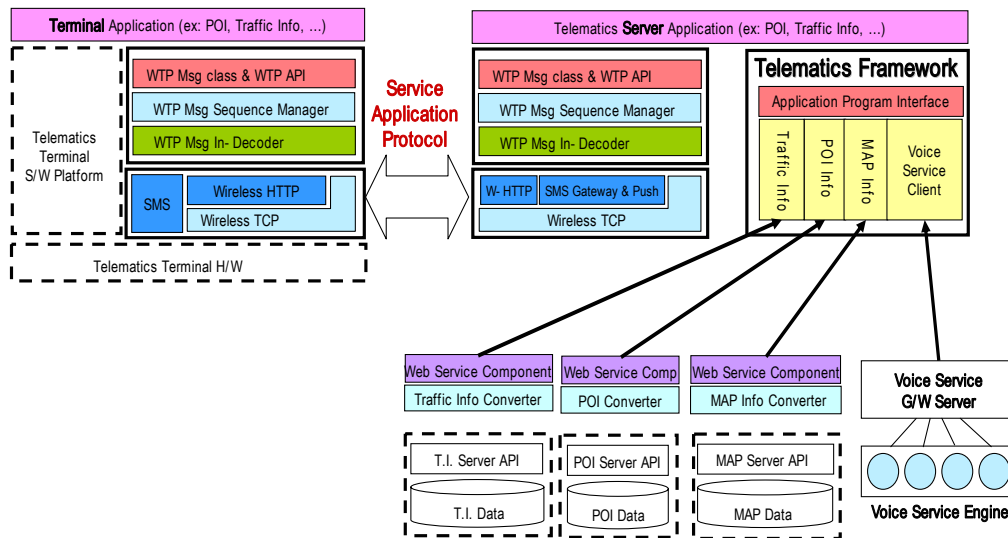


Fig.3. Framework based telematics system

Fig. 3 shows a telematics system based on telematics framework. Telematics framework is located in telematics service provider system providing application programming interface to communicate with content provider system and to get useful contents for telematics services. Traffic information system, Point-of-Interest data system, map data system and voice service center can be located in any place on the internet. Web service components for each content provider are placed between content provider and telematics service application; however they could be in the service provider system or in the content provider system.

In this study, we consider the voice service center as a content provider which receives requests to convert speech to text or text to speech and returns processed results, in other words, text and speech are kinds of contents to build a telematics service. Therefore we put some APIs in the telematics framework to communicate with voice service center which can process voice related converting requests.

### 3.2 Telematics service provider

A telematics service provider server makes connection with telematics terminals on mobile networks providing users with services using telematics gateway. And it communicates with external content providers to gain useful data for services using telematics framework.

Voice mail server and Traffic & POI information server run separately for each service. Each application server uses various common components such as terminal manager for using and managing features of various kinds of telematics terminals.

Content provider manager uses and manages external content servers where service provider gains actual contents from. VoiceXML processor analyzes voiceXML documents and makes decisions what to do next. Each service processor manages each service and makes decision upon reply from users.

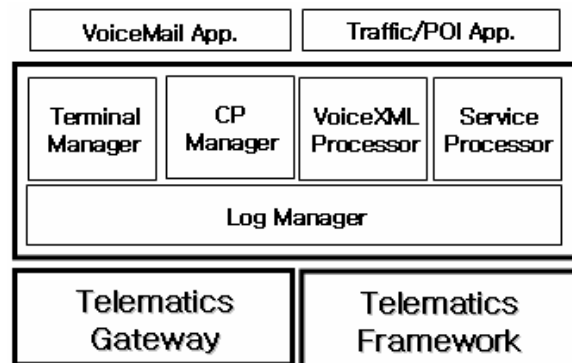


Fig.4. Service Provider Server

### 3.3 Service Scenario Server

Service scenario server processes requests form a service provider server. The server produces VoiceXML documents in reply, which are processed by the service provider server. It plays a role as a document server in VoiceXML specification.

The scenario server takes email data, traffic & POI data and service related data from TSP server and produces voiceXML documents which are the next scenarios of the services using VoiceXML manager and returns them to TSP server. The server manages TSP servers to communicate using TSP manager and stores and processes data and features of services using manager of each service. The server uses HTTP

protocol to receive requests from TSP server and reply to it. And it uses a database to store and manage a lot of data from TSP server.

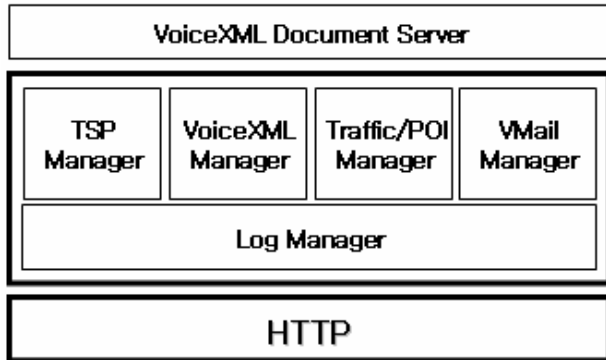


Fig.5. Service Scenario Server

### 3.4 Voice Service Center

A voice service center provides services converting text to speech and speech to text using a voice recognition engine and a voice synthesis engine. The center produces all needed voices that TSP server provides to service users and converts user requests in voice form to ones in text form.

When a user in a car on road gives a command with his voice, TSP server receives the voice and asks a voice service center to recognize and convert it to corresponding text. Then the voice service center use voice recognition engine to convert it. If it fails to recognize the voice, the center sends the voice to the human operator. The human operator listens to the voice and input corresponding text. In this way, voice service center always returns an exact result to TSP server. It means that a voice based service provided by TSP server does not disturb driver to speak several times to make an acceptable command, which helps him focus on his driving safety.

### 4. System Implementation

We developed two telematics services. One is a traffic & POI information service in which a user gives command with his voice and gets useful information about traffic and shops. The service shows traffic information and required shop information on a map based on the location of the car. TSP server gets traffic information used in this service from TELIC(Telematics Information Center) which is independent information center on the Internet in Korea using telematics framework technology.

The other is a voice mail service in which a user requests to give a list of new mails and the contents of the mails with his voice, and the system reads a list of new mails or contents of them to the user showing corresponding text on the screen as shown Fig.6.



Fig.6. Voice based telematics services

### 5. Conclusion

Telematics services are really useful, mainly because of the proliferation of automobile, telecommunication devices and computers which allow access to various services at any places such as moving cars. However, voice based telematics services must be provided to ensure that drivers keep their eyes on the road and hands on the wheel for their safety.

In this paper we have presented a system which allows providing voice based telematics services. The system is based on an open telematics system solution which enables service providers to build their services using external contents, and we showed that voice conversion could be content for telematics services. The voice conversion is done in server side with voice recognition engine and an human operator not in a telematics terminal to guarantee high quality of voice processing service. The voice based service system uses VoiceXML technology to be accommodated to service changes easily

*References:*

- [1] W. Y. Han, O. C. Kwon, J. H. Park & J. Kang, A Gateway and Framework for Telematics Systems Independent on Mobile Networks, ETRI Journal, Vol. 27. No. 1, Feb. 2005, p106-109
- [2] 2. Stein, A.C., Parseghian, Z., and Allen, R.W. (1987). A simulator study of the safety implications of cellular mobile phone use (Paper No. 405). Hawthorne, CA: Systems Technology, Inc.
- [3] 3. Zwahlen, H.T., Adams, C.C., Jr., and Schwartz, P.J. (1988). Safety aspects of cellular telephones in automobiles (Paper No. 88058). Proceedings of the ISATA Conference, Florence, Italy.
- [4] 4. Wickens, C.D. (1984) Processing resources in attention. In R. Parasuraman and D.R. Davies (Eds.), Varieties of Attention. London: Academic Press.
- [5] 5. McMillan, G.R. and Eggleston, R.G. (1997). Nonconventional controls. In G. Salvendy (Ed.), Handbook of Human Factors and Ergonomics. New York: John Wiley & Sons, Inc.
- [6] 6. Wargnier, J. (2002). Considerations for robust speech recognition and sound quality for automotive handsfree kits. 21st Annual Conference of the Applied Voice Input/Output Society (AVIOS) pp.339-348
- [7] 7. Asatani, K. (1998). Standardization on Multimedia Communications: Computer-Telephony-Integration-Related Issues, IEEE Communications Magazine pp.105-109