Applying Mobile Agent for Internet-based Distributed Speech Recognition

Emrul Hamide Md Saaim, Mohamad Ashari Alias, Abdul Manan Ahmad, Jamal Nasir Ahmad
Speech Recognition Laboratory,
Faculty of Computer Science and Information Systems,
University of Technology Malaysia
81310, Skudai, Johore, Malaysia
Tel: +60-7-553-2201, Fax: +60-7-556-5044

Abstract

There are several application have been developed on internet-based speech recognition. Internet-based speech recognition is a distributed application and there were various techniques and methods have been using for that purposed. Currently, client-server paradigm was one of the popular technique that been using for client-server communication in web application. However, there is a new paradigm with the same purpose: mobile agent technology. Mobile agent technology has several advantages working on distributed internet-based system. This paper presents, applying mobile agent technology in internet-based speech recognition which based on client-server processing architecture.

1.0   Introduction

Motivated by the explosive growth of the internet, both speech researcher and software engineer have been working on the integration on speech technologies into Internet application [2]. Generally, there are three alternative strategies in the design of internet-based distributed speech recognition architecture [3]. The first is server-only processing in which all processing are done at the server side and the speech signal is transmitted to the server either trough the internet by using speech coding or via a second channel like telephone [2,4]. The secondly is client-only processing in which most of speech processing is done at the client-side and then the results are transmits to the server [5]. Third is client-server processing. In this model, front-end processing is done at the client-side. Then the speech features are transmitted to the server and finally the processing of speech decoding and language understanding are performs at the server-side. However, all the distributed speech recognition strategies were based on client-server model.

The aim of this paper is to apply mobile agent technology in internet-based distributed speech recognition. One of the most attractive applications for mobile agents is distributed application, particularly in mobile-computing scenarios. By moving the code to the data, a mobile agent can reduce the latency of individual steps, avoid network transmission of intermediate data, continue work even the presence of network disconnections, and complete the overall task much faster than a traditional client-server solution [10]. Based on those advantages, we propose a new architecture in applying mobile agent technology for internet-based speech recognition.

2.0   Mobile Agent

Mobile agents are programs that can move through a network under their own control, migrating from host to host and interacting with other agents and resources on each. A concept

of mobile agents always been applied in a network computing such as distributed system, information retrieval, network management or other else.

This paper has been focus on distributed system, which our domain is speech recognition system and we try to apply mobile agent technology for that purposes. The mobile agent technology promotes a new paradigm for distributed computing which is particularly suited to the design and implementation of internet applications [6]. One more thing, mobile agent are an effective paradigm for distributed applications, and are particularly attractive for partially connected computing [11]. Mobile agents have been developed as an extension to mobile code approach (e.g. applet) and could replace the client-server model and its architecture in the near future [9].

The general role of mobile agents in internet applications is to (i) proactively move through a multiplicity of different execution frameworks, and (ii) there interact with local resources and other mobile agents [6]. Beside that, compare to traditional distributed computing schemes, mobile agents promise (at least in many cases) to cope efficiently and elegantly with a dynamic, heterogeneous and open environment which is characteristic for today's internet [7].

This approach has several benefits [1]:

- Reduced complexity for user and agents. Requester need only specify the work to be done and advice on its execution. Agents then focus on performing their specialized task, not on coordinating execution and result.
- Participating a more open and dynamically extensible computing style wherein agents writing in many language and styles can work together. New or different agents can be added or replaced on the fly without requiring reprogramming to take advantage of their capabilities.
- Encouraging reuse across applications and domains because inter-agent interactions are not pre-defined and their interfaces are not hard-coded.

Another thing about applying mobile agent technology in distributed applications, mobile agent also ease the development, testing and deployment of distributed applications since they hide the communication channels but not the location of the computation [11]. They eliminate the need to detect and handle network failure except during migration. They do not require the pre-installation of application-specific software at each site and they can dynamically distribute and redistribute themselves throughout the network.

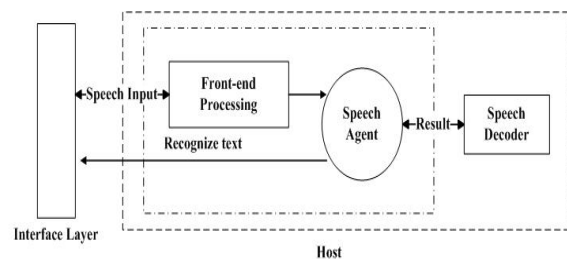## 3.0 The Architecture Design

Figure 3: Design of Speech Recognition Agent Architecture

Generally, the design of speech recognition agent architecture is consists of three main parts. There is interface layer, speech recognition system, which divided into two part, front-end processing and speech decoder and lastly, speech agent. The architecture of the system is shown in figure 1. At the interface layer site, there is the web browser, the applet and the voice recording process. On the host site, there is the web server, the speech agent, the front-end processing and the speech decoder.

Interface layer is responsible to get speech input and displaying the recognized text. However, all tasks will be cover up by web browser. Web browser is responsible for recording the speech voice and transmitting the speech data to the applet. The applet will transmit the speech data to the front-end processing. All steps in front-end processing such as acoustic processing and features extraction will be process before the result have to transmit to the speech agent. Speech agent is a part of mobile agent concept. Main function of speech agent is to control transmission, receiving speech data form front-end processing, speech decoder, and vice-versa. The speech agent also allow to process speech recognition in asynchronous method, which it is an advantage for using mobile agent concept. Therefore, speech agents will handle every request from remote computer safely. Furthermore, they also can reduce the network traffics. Speech decoder is responsible to recognize the speech data and transmit the recognized text to speech agent again. Finally, speech agent will transmit the outcome to the interface layer for displaying the recognized text to the user.

4.0 Speech Recognition Agent Framework Model

The whole of speech recognition agent framework model in detail is shown in figure 4. It contains several main component, interface layer, Sphinx-4 front-end processing, Sphinx-4 decoder, master-agent and slave-agent. The description of those components can be found in the succeeding portion of the paper.
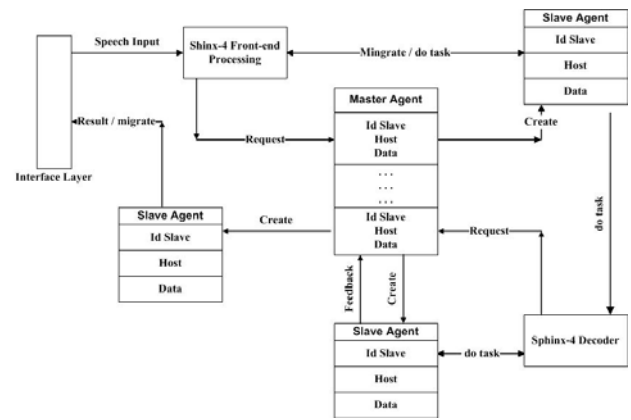

Figure 2: The Architecture Design of Speech Recognition System

4.1 Interface Layer

Interface layer is responsible to record voice and as an interface between system and user. Interface layer is a web application (JSP). The basic task for the interface layer is record user voice and sends to the Sphinx-4 front-end processing. After the users recorded theirs voice, they can submit their voice from the interface layer.

4.2 Sphinx-4 Front-end Processing and Speech Decoder

Sphinx-4 Front-end processing is one of the speech recognition components. Sphinx-4 Front-end processing is responsible to extract data from voice recorded. The Sphinx-4 front-end processing contains feature extractions and acoustic processing. After the acoustic processing and features extraction has been done, speech encoder will request to master-agent. Master-agent will send a slave-agent to take the result from speech encoder.

Sphinx-4 speech decoder will recognize the speech data. Sphinx-4 speech decoder will received the speech data from master-agent through slave-agent. After sphinx-4 speech decoder recognized the speech data, it will

request to the master-agent to take the recognized text.

In this paper, both front-end processing and speech recognizer were based on Sphinx-4 architectures. Sphinx-4 is a state-of-the-art speech recognition system written entirely in the Java^TM programming language.

4.3   Speech Agent

Speech agent is responsible to receive the speech data that had been encoded by the Sphinx-4 front-end processing and send back the output (in text format) that had been recognized by the Sphinx-4 decoder to the interface layer.

There are a few techniques in mobile agent concepts, such as master-slave pattern. Based on figure 2, speech agent using master-slave technique to receive and transmit the speech data and recognize text.

In master-slave patterns, speech-agent will divide into two part, master-agent and slave-agent and both of them have their own task. Master-agent is responsible to monitoring the process between slave-agent and Sphinx-4 front-end processing, Sphinx-4 decoder. Master-agent will monitor the transmitted and received data between them, such as the name of the host, what task they should do and assigned every slave-agent with their own identification.

Slave-agent is responsible to performs it tasks that given by master-agent. For examples, ones the slave-agent get the assigned task, they moves to Sphinx-4 front-end processing, Sphinx-4 decoder or interface layer to performs the task. After that, the slave-agent will returns with the result of the task to the master-agent. Slave-agent contains several items such as their identification, the host identification and the assigned task.

5.0   Conclusion

Applying mobile agent in internet-based distributed speech recognition was presented. Mobile agent with master-slave pattern design approach will be applied to the internet-based distributed speech recognition to support monitoring that distributed systems. With the advantages for using mobile agent for web application such as reduce the network load, robust and fault-tolerant, execute asynchronously and autonomously and other else. Therefore, with these advantages, mobile agents are used to support distributed speech recognition to reduce the complexity of the systems.

Reference

[1]   David Martin et. al, "Open Agent Architecture", http://www.i.sri.com/~oaa

[2]   Zhemin Thu and Philipos C. Loizou, (1999), "Speech Recognition Over The Internet Using Java", *International Conference on Acoutics, Speech and Signal Processing, Phoenix, AZ:2367-70.*

[3]   Wei Qi Zhang et. al, (2000), "The Study on Distributed Speech Recognition System"

[4]   Domonic Vaufreydaz et. al, (1999), "A network architecture for building applications that use speech recognition and/or synthesis", *European Conference an speech Communication and Technology, pages 2159-62.*

[5]   Aldebaro Klautau, (2000), "Server-assisted speech recognition over the internet", *The 2000 IEEE International Conference on Acoustics, Speech and Signal Processing.*

[6]   Marco Cremonici et. al, (1999), "Building Mobile Agent Applications in HiMAT", *IEEE International Symposium on Software Engineering for Parallel and Distributed*

Systems (PDSE '99), Los Angeles (CA), USA.

[7]   Stefan Funfocken and Friedemann Mattern, (1999), "Mobile Agents as an Architectural Concept for Internet-based Distributed Applications - The WASP Project approach", *In: Steinmetz (Ed): Proc. KIVS'99, pp. 32-43*

[8]   Timon C. Du et. Al, (2003), "Mobile Agents in Distributed Network Management", *Communication of the ACM, Vol.46. No.7*

[9]   Aneiba, A and Rees, J.S., (2004), "*Mobile Agent Technology and Mobility"*, *Proceeding of the 5th Annual Postgraduate Symposium on the Convergence of Telecommunications, Networking and Broadcasting, pp. 14-20,PGNet 2004, Liverpool, ISBN:1-9025-6010-8*

[10]   Robert S. Gray et. Al, (2001), "Mobile Agents versus Client/Server Performance: Scalability in an Information-Retrieval Task", *In Proceeding of Mobile Agents 2001*

[11]   Robert Gray et. Al, (1996), "Mobile Agent for Mobile Computing", *Technical Report PCS-TR96-285.*