

Cross-Language Instant Messaging

Tippu Junaidi*, Umesh Kamble*, Viraj Kalkur*, Haridas Kavindra*, Y.C.Kiran **

* Dayananada Sagar College of Engineering, Bangalore

** Associate Professor, Department of Information Science and Engineering, Dayananada Sagar College of Engineering, Bangalore

Abstract- Along with the rapid development of Internet technologies, the instant messaging has become nowadays an important medium for a huge number of people to communicate with friends, family, and even colleagues while working. People who come from different corner of the world speak different native- languages. Even if the instant messaging technology is now so developed, there is a barrier to communication for people having different native-languages.

This application tries to let people easily chat without having to be familiar with the others' native language by integrating "instant messaging" and "machine translation" technologies. This shall overcome the language barrier to communication. Furthermore, with proper design, this mechanism can also facilitate conversation practice.

I. INTRODUCTION

In recent years, along with the development of Internet communication technologies, various network-related applications are springing up. Social network and its related applications are the hottest topics. Among them, the instant messaging (IM) has become nowadays an important medium for people to communicate for its convenience and free of use. Through the Internet we are able to make friends with people around the world and chat with them using a computer. The instant messaging has shortened the geographical distance between people all over the world - the conversation is as easy as sitting in front of the computer and popping fingers to type – the text communication has become easy and efficient. However, the invisible distance – the barrier results from the different native-languages people speak has not been eliminated yet. Problems are with inputting foreign languages and to even understand other languages. It is still a tough work for people to communicate with each other's while they are speaking different native languages.

This language barrier needs to be overcome by the technology of natural language processing (NLP). Natural language processing Technology, such as information retrieval, speech recognition, machine translation, automatic summary and so on, has developed rapidly in recent years. The machine translation technology can be used to do the language translation task instead of human translators or language experts. If "instant messaging" and "machine translation" technologies cooperate so that instant messaging is no longer just a messenger's role: it also properly does real-time translation to the context of messages. This shall eliminate the barrier to communication easily for those whom are speaking various different languages. Furthermore, with some proper design, this mechanism can also provide some facilitation for conversation practice during language learning.

II. RELATED WORKS

2.1 Instant messaging

Instant messaging is a kind of network service that allows two or more people to make text chat to each other's. It is developing rapidly in recent years, and is integrated with more and more functions such as offline message delivery, voice chat, video chat, file-transfer etc. Instant messaging is now no doubt one of the most popular network services in the Internet.

In April 2009, "comScore, Inc." released a result [1] of an application Internet users in France revealed that people in France spent the highest share of total time spent at 14.3 per cent on instant messaging, followed by social networking at 5.7 per cent. In October 2009, another report of "comScore" [2] indicated that online communications, entertainment and social networking occupied the highest share of Hong Kong Internet users' attention. Instant Messengers accounted for the highest share of minutes spent online at 16 per cent. Yet another survey [3] indicated that MSN Messenger [4] has the strongest penetration worldwide, with 61 per cent of worldwide IM users utilizing the application. MSN Messenger is also dominant in Latin America, reaching more than 90 per cent of IM users, and in Europe and Asia Pacific, reaching more than 70 per cent of IM users in each region.

North America is the most competitive IM market, with MSN Messenger, AOL/Aim [5] and Yahoo! Messenger [6] each garnering between 27 per cent and 37 per cent of IM users.

According to a survey in Taiwan by InsightXplorer Limited [7], the current usage rate of instant messaging in Taiwan is 92.7%, among which 50.1% of netizens have installed two or more instant messaging software. From the view of user-age, the younger the netizens are, the more often they communicate with friends via instant messaging. In the ages of 15 to 19 year-old, 99% Web surfers use instant messaging; despite from the teenagers, in the age of 35 to 39 year-old, the usage rate is as high as 81.3%. As for the user population, MSN Messenger owns a number of more than 700 million users, while the Yahoo! Messenger about 470 million users.

2.2 Electronic dictionary

The software electronic dictionary is usually available in the form of either package software or network service accessed through a Web browser (online service). For the latter, because of its no installation of software needed and free to use, is currently most popular way to provide the dictionary service, commons are Yahoo! dictionary [8], Google translate [9], Microsoft Bing translations [10]. These online dictionary services normally provide the functions of vocabulary queries, paragraph translation, webpage translation, text file translation and some others. Google translate and Microsoft

Bing translations provide not only a number of intimate small tools and resources to end-user, but also application programming interface (API) to the computer software developer [11, 12]. One example of benefit from automated translation technology is that many Internet Content Providers make copies of the original web content available in various language versions generated by machine translation, thereby expanding the scope of audience to global.

2.3 Instant Messaging with Language Translation

With today's technology, the cross-language dialogue is not out of reach. The key is to cooperate "instant messaging" and "electronic dictionary" technologies. For Internet instant messaging, not only passing messages for the two sides in dialogue, but also makes proper language translation to the context of the messages. The electronic dictionary can be used to facilitate the translation task.

Instant messaging is a kind of network service that two or more users can talk to each other's by real-time text chat. Among many, MSN Messenger [4] is developing quickly and is one of the most popular globally. As for the electronic dictionary, now often provides service through web pages (also referred to as Internet dictionary, online dictionary). In addition to look up words, it also provides advanced sentence translation function. Among many, Google Translate [9] is one of the most developed online dictionaries and it comes with application programming interface (API) that can be used by computer program rather than human [11,12].

This developed application uses an instant Messenger for communication technology and Google Translate for language translation technology, to design a system that supports "bi-directional multi-language translation" instant messaging, as shown in Fig. 1.

message context according to the native language settings of two ends.

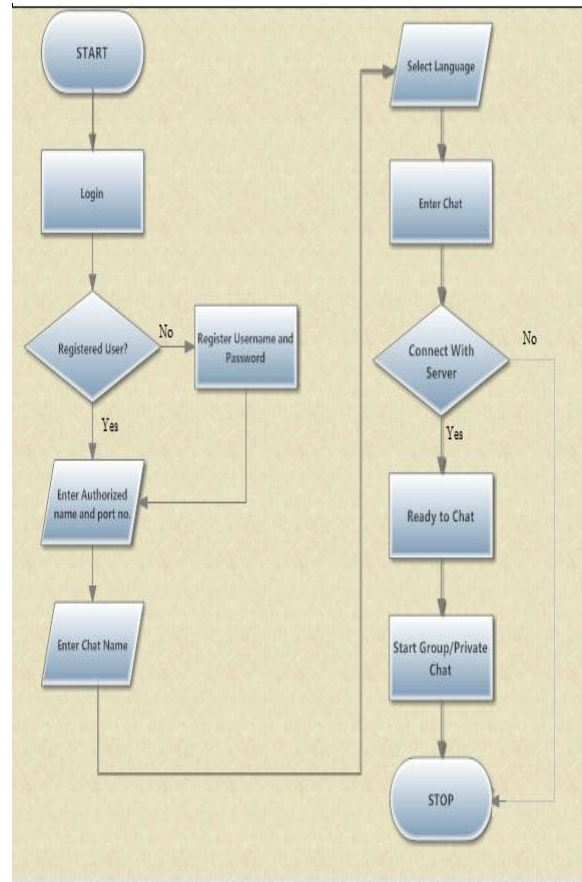


Fig: Flowchart for Client Module

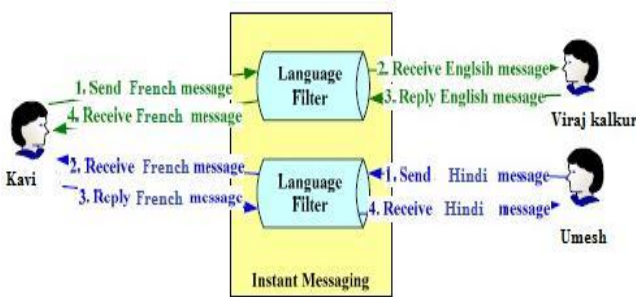


Fig1: Bi-directional multi-language automatic translation instant messaging.

III. SYSTEM PROCESS AND ARCHITECTURE

3.1 System Process

The system provides a set of management interfaces to allow users to maintain their MSN Messenger contact list, each contact in the list can be configured to be associated with a value of "native-language". This setting will be stored in a database for future reference. The system is responsible for passing text messages between two dialogue ends, and more importantly, making appropriate language translation to the

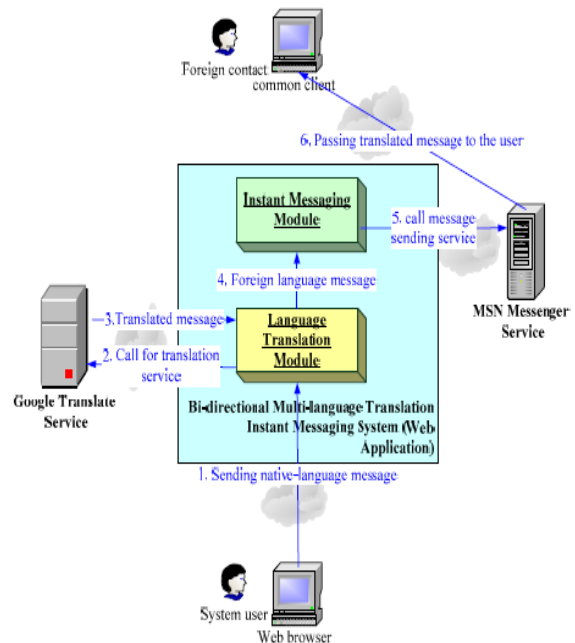


Fig2: Sending Process by the instant messaging system.

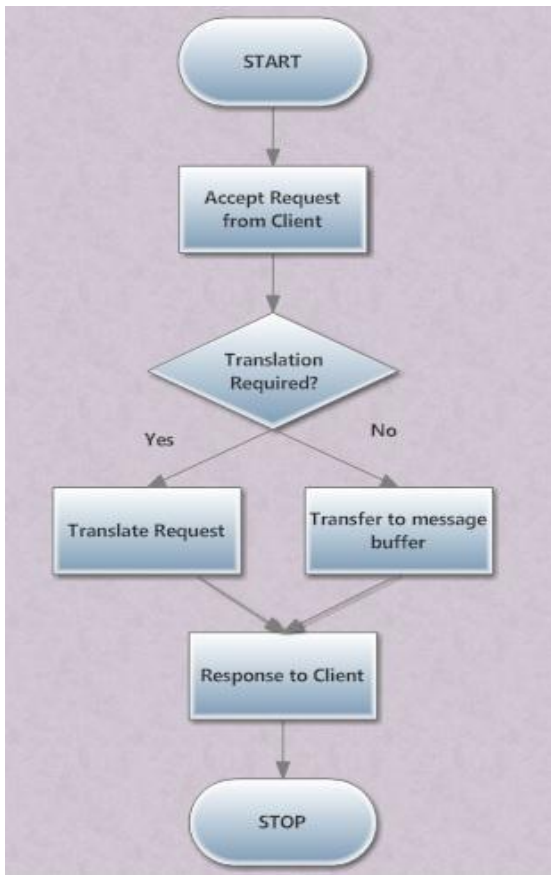


Fig: Flowchart for Server Module.

messenger client (which end is denoted as "common client" in fig. 2 and fig. 3). Because of the nature of the system design, the end using common client even does not need to know the existence of this system which helps with the translation. This makes it more practical and feasible of the system.

3.2 System Architecture

The system aims at integrating the online instant messaging service - Messenger with bi- directional multi-language translation capabilities, allowing users to easily chat with friends around the world using their own native language, and can from which to achieve language learning.

The tools and development environment used for building the Cross language messenger are as follows:

- i. Programming language - Java SE
- ii. Integrated Development Environment (IDE) - NetBeans IDE
- iii. Database- MySQL Workbench [14]
- iv. Web Server - Apache Tomcat [15]
- v. Instant messaging API - Java MSN Messenger Library (JML) [16]
- vi. Language translation API - Google Translate API for Java [12]

System Architecture

System architecture is as follows:

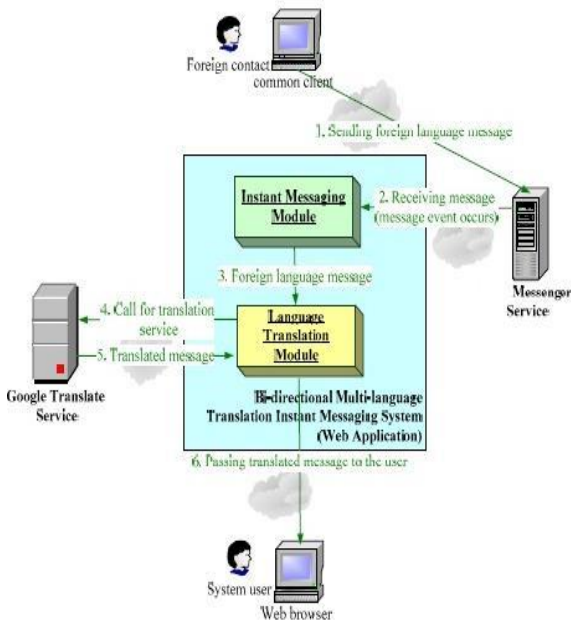
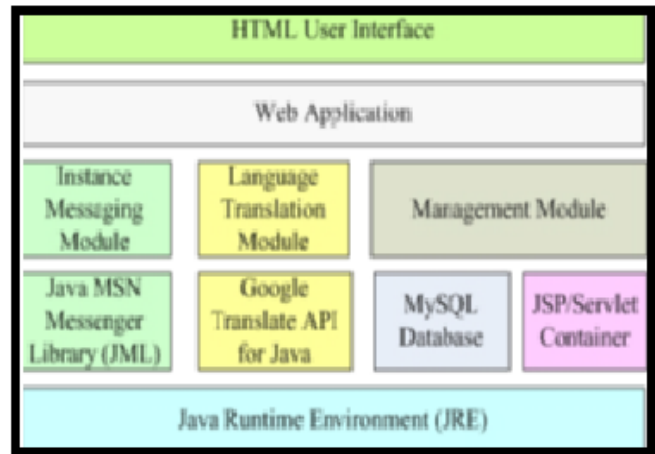


Fig.3: Receiving process of the instant messaging by system.

Fig. 2 and fig. 3 show that in order to achieve bilingual conversion instant messaging ability, only one of the two ends of user needs to chat through the system (which end is denoted as "user" in fig. 2 and fig. 3), the other end of user can use common

Instant Messaging Module: mainly responsible for account login and authentication, access to contact lists, message transmission and reception. The functions of this module will be realized with the JML APIs.

Language Translation Module: responsible for natural language translation capable of bi- directional multi-language translation ability. For example, to use French-English translation, entering an English sentence to the module then the module will translate the context into English and output to the caller. This module will cooperate with the Google Translate API for Java to realize the functions.

Management Module: The management module is responsible to record/retrieve the dialogue history and setup native-language configuration.

User Interface: responsible for integrating and coordinating the two modules: instant messaging and language translation, and provides user interface.

IV. SYSTEM IMPLEMENTATION

An enhanced version of the Internet instant messaging system has been designed with the following capabilities: Every contact in the messaging contact list can be set to associate with a language option which denotes the native language of the contact.

User chats with contacts who speak different native language from the user without having to be familiar with each other's native languages. The system translates the message context into the proper languages according to the language options associated to the contacts, and then convey the messages to them.

V. CONCLUSION

The advance of the Internet technology has eliminated the barriers of the geographic distance. However, with the perspective of the global village, the barrier to communication results from various different native languages people speak is still a problem to be solved at present.

Through the design of this work that cooperates instant messaging and machine translation technologies, having simultaneous interpreter is no longer of the privilege of a country leader, it is now possible for everyone to take to enjoy the convenience of cross- language communication online.

Wide range of situations can benefit from this, from making foreign friends to the cross- language global customer service. The additional foreign language conversation learning function is also a good helper.

REFERENCES

- [1] "Instant messaging most popular online activity in France", comScore Inc. Press Release, April 6, 2009 http://www.comscore.com/Press_Events/Press_Releases/2009/4/Instant_Messaging_Most_Popular_Online_Activity_in_France.
- [2] "Hong Kong Internet users spend twice as much time on instant messengers as counterparts in Asia-Pacific region", comScore Inc. Press Release, October 2009 <http://ir.comscore.com/releasedetail.cfm?ReleaseID=415641>
- [3] "Europe surpasses north America in instant messenger users", comScore Inc. Press Release, 2006 http://www.comscore.com/Press_Events/Press_Releases/2006/04/Europe_Surpasses_North_America_in_Instant_Messenger_Usage
- [4] Google AJAX Language API, <http://code.google.com/intl/zh-TW/apis/ajaxlanguage/> [12] Google Translate API for Java, <http://code.google.com/p/google-api-translate-java/>
- [5] NetBeans IDE, <http://netbeans.org/>
- [6] MySQL Community Server <http://www.mysql.com/>
- [7] "Instant messaging population up to 90 percent" Insightxplorer Inc. Press Release, March 2006, http://www.insightxplorer.com/news/news_03_24_06.html
- [8] Java MSN Messenger Library (JML)

AUTHORS

First Author – Tippu Junaidi -Bachelor of Engineering (B.E), Dayananada Sagar College of Engineering, Bangalore. tippujunaidi@gmail.com (+91 9535042422)

Second Author – Umesh Kamble Bachelor of Engineering (B.E), Dayananada Sagar College of Engineering, Bangalore. Umi.kamble@gmail.com (+91 9035025805)

Third Author – Viraj Kalkur- Bachelor of Engineering (B.E), Dayananada Sagar College of Engineering, Bangalore. virajkalkur@gmail.com

Third Author – Haridas Kavindra- Bachelor of Engineering (B.E), Dayananada Sagar College of Engineering, Bangalore. Kvi91@yahoo.fr