Performance Analysis of a P2P-Based VoIP Software

Gao Lisha and Luo Junzhou
Department of Computer Science and Engineering
Southeast University, Nanjing, China
{lsgao, jluo}@seu.edu.cn

Abstract

With the development of network, multimedia will be the main application in next generation network. Voice is one of the most important applications. Recently a kind of P2P-based VoIP software, Skype, has been receiving more and more attention both in academia and industry. Skype claims that it's better than other VoIP software, because of its high call completion rate and superior sound quality. This paper will reveal Skype's technique and have a performance comparison between Skype and MSN Messenger, which uses traditional VoIP protocol. The result indicates that the voice quality of Skype is no better than traditional VoIP software and the great benefit of P2P involved with VoIP is that it can solve NAT and firewall problems.

1. Introduction

Recently a new technique combining VoIP and P2P (peer-to-peer) appears. Skype is such a software developed by KaZaa in 2003. Although there are many IP telephone software: MSN Messenger, Net2Phone, ICQ .etc and they all have distinct features, yet all of them use traditional VoIP protocol such as SIP and H.323 [1, 2, 3]. Skype is widely considered to be the most popular IP telephone for its high call completion rate and superior sound quality.

With their experience in creating the most popular decentralized P2P network (KaZaA) and the largest P2P-based content distribution network (Joltid PeerEnabler), the Skype team has succeeded in leveraging all of the resources available in the network [4]. This has raised the call completion rate and quality in the Skype network. And it is all achieved without the need for costly centralized resources.

According to the Skype official website, Skype is a free program that uses the latest P2P technology to bring affordable and high quality voice communications to people all over the world [5]. Skype is similar to the MSN Messenger and Yahoo IM applications, as it has capabilities for voice-calls, instant messaging, audio conferencing, and buddy lists. Skype users can speak to other Skype users for free, call traditional telephone numbers for a fee, receive calls from traditional phones, and receive voicemail messages. However, the underlying protocols it employs are quite different.

The main difference between Skype and other VoIP clients is that it's a P2P overlay network rather than the traditional server-client model. The Skype user directory is entirely decentralised and distributed among the nodes in the network, which means the network can scale very easily to large sizes without a complex and costly centralised infrastructure. They utilize the processing and networking power of the end-users machines. Each new node added to the network adds potential processing power and bandwidth to the network [4].

The remainder of the paper is organized as follows. Key components of the Skype software and the Skype network will be presented in section 2. In section 3, we'll set up an experiment to have a performance comparison between Skype and MSN Messenger. Finally we'll give a conclusion.

2. Technical aspects about Skype

This section will discuss some technical aspects about Skype in detail. Some has been explained in Skype website [4] and its forum [5].

2.1. Main working flow

With P2P, a user could search a directory using a special browser and then click on the name he or she wanted to call. The client software would send the request to a super-node that would search a giant routing table listing users and their current IP addresses, and send back the result to the client so the VoIP call could be set up, as shown in Figure 1. Then the communication will take place between the two users, not via the super-node.

Proceedings of the Advanced International Conference on Telecommunications and
International Conference on Internet and Web Applications and Services (AICT/ICIW 2006)
0-7695-2522-9/06 $20.00 © 2006 **IEEE**

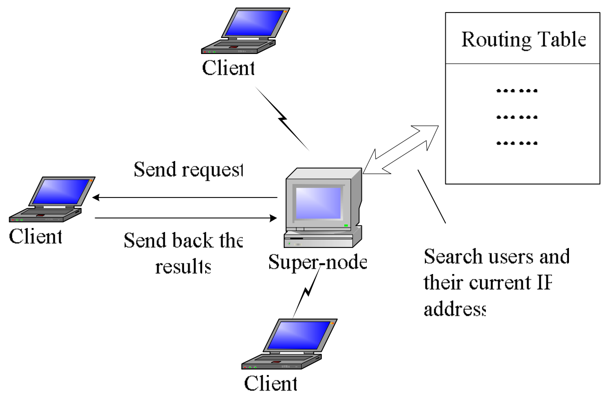

Figure 1. Main working flow of Skype

The super-node functionality is built into the client; if a powerful computer with a fast network connection runs the client software, it will automatically become a super-node, effectively acting as a temporary indexing server for other, slower clients.

This is pretty much how file sharing works, and file sharing can easily scale to the millions of users that might take advantage of such a system.

2.2. Global decentralized user directory

Most instant message or communication tools track users with a central directory which logs each username and IP number and keeps track of whether users are online or not. But central directories are extremely costly when the user base scales into the millions.

For Skype, there is a centralized database where users are registered and issued a password and membership certificate, but all the traffic is decentralized. In order to deliver high quality telephony with the lowest possible costs, Skype uses a third generation of P2P technology ("3G P2P"): Global Index (GI) technology. The Global Index technology is a multi-tiered network where super-nodes communicate in the way that every node in the network has full knowledge of all available users and resources with minimal latency [4].

2.3. Host cache

The Skype network is an overlay network, thus each Skype client should build and refresh a table of reachable nodes called host cache. It contains IP address and port number of super-nodes and refreshes regularly. It is the most critical part in the Skype operation. At least one valid entry must be presented in host cache. A valid entry is an IP address and port number of an online Skype node [6].

2.4. Bandwidth

The bandwidth used for a call varies depending on several factors like user's Internet connection, computer processing power. Skype tries to optimize bandwidth usage so that people with less bandwidth can use Skype, while at the same time people with powerful computers and high-bandwidth connections can get the better voice quality that high bandwidth allows. Most of the time, on a broadband connection, Skype calls use 3~16 kb/s in a call, and 0~0.5 kb/s while idle.

2.5. CODEC

Skype uses the GIPS iSAC Codec. Skype has partnered with Global IP Sound, the telephony components of the software including echo cancellation and various codecs that translate analog voice into digital packets. GIPS iSAC is a wideband, adaptive Codec, designed to deliver high quality sound in both high-bit-rate and low-bit-rate conditions [7]. GIPS iSAC makes VoIP communications possible even using a dial-up modem and automatically adjusts transmission rates to deliver better-than-PSTN voice quality.

2.6. Firewall and NAT traversal

VoIP is one of many IP applications that have problems traversing NATs and firewalls [1]. Most Voice-over-IP applications don't work from behind firewalls and NAT (Network Address Translation) devices. But nearly all broadband users are behind a NAT or firewall and so they cannot use VoIP applications.

Skype routes calls through multiple peers, to get around the firewall problem. For example, if there are two people both behind firewalls, the software routes the call to a third person who is not behind a firewall and both connect to this person And the call is routed via that peer. Thus the two people set up the connection.

2.7. Intelligent routing

Skype is able to intelligently route encrypted calls through the most effective path by using every possible resource. Skype even keeps multiple connection paths open and dynamically chooses the one best suited at the time. This has the great effect in reducing latency and increasing call quality throughout the network [4].

2.8. Security

Proceedings of the Advanced International Conference on Telecommunications and
International Conference on Internet and Web Applications and Services (AICT/ICIW 2006)
0-7695-2522-9/06 $20.00 © 2006 IEEE

Skype encrypts all calls and instant messages end-to-end for privacy. Skype is encrypted end-to-end because it uses the public Internet to transport voice calls and text messages and sometimes these calls are routed through other peers. Skype encryption ensures that no other party can eavesdrop on users' call or read instant messages. AES (Advanced Encryption Standard) is used in Skype to protect sensitive information. Skype uses 256-bit encryption, which has a total of 1.1 x 1077 possible keys, in order to actively encrypt the data in each Skype call or instant message. Skype uses 1536 to 2048 bit RSA to negotiate symmetric AES keys. User public keys are certified by Skype server at login.

3. Experiment and analyzing

Many factors determine voice quality, including the choice of codec, echo control, packet loss, delay, jitter, and the design of the network [1]. In order to compare the performance of VoIP software that use P2P technique and other protocol (such as SIP and H.323), we designed the experiment scheme. We selected the widely used instant message software - MSN Messenger developed by Microsoft to compare with Skype.

MSN Messenger is a kind of communication software, which supports instant message, audio conversation, web cam and other functions. It uses SIP (Session Initial Protocol) as VoIP protocol. Its voice quality is good, but it can't work through firework and NAT. So people often can't setup connection.

3.1 Experiment setup

All experiments were performed for Skype version 0.97.0.6. Skype was installed on two Windows XP machines. One machine was a Pentium II 200MHz with 128 MB RAM, and the other machine was a Pentium Pro 200 MHz with 128 MB RAM. Each machine had a 10/100 Mb/s Ethernet card and was connected to a 100 Mb/s network.

We performed experiments under different kinds of network scenarios, including: LAN (same/different domain), WLAN (same /different domain), DSL (BJ Telecom), GPRS (China Mobile), CDMA2000x (China Unicom). We test parameters in MSRA lab for most scenarios, as we own the basic devices, such as GPRS, wireless card, etc. So if we say "in same domain", we test the parameters both in MSRA Domain (domain name: msrchina.research.microsoft.com); and if we say "in different domains", it means we use one PC in MSRA Domain, and the other in SEU Domain (domain name: seu.edu.cn).

All experiments were performed between February and March, 2004.

3.2 Result and Analysis

We measure each performance of Skype and MSN messenger separately according to [8, 9, 10]. Our parameters include presence, connection setup delay, M2E delay, MOS and handover support. To shown figure compendiously, we use *W* denote WLAN, *C* denotes CDMA, *L* denotes LAN, *G* denotes GPRS, *D* denotes DSL. '*/s*' means that the test scenarios are in same domain, and '*/d*' means they are in different domains. Due to the instability of software and network condition, some test results will fluctuate in a certain bound. So these parameters in figures are showed in a bar to indicate their fluctuant bounds.

3.2.1. Presence. Presence indicates the refresh rate of the software. If a friend has logged in/out, it denotes how soon the panel will refresh his status. We use *second* as measure unit and measure the two kinds of software respectively in various scenarios. The results are given in Figure 2 and Figure 3. Since there is much difference between Skype and MSN Messenger, it can't be seen clearly in one figure. We zoom in the figure that shows the presence of MSN Messenger.

From Figure 2 and Figure 3, we can see the presence of Skype is very poor. It always doesn't show the correct status. Usually, when a user has been online over 5 minutes, even when both sides begin to communication, Skype still shows that the person is off-line.

By contrast, MSN Messenger does better. The refresh rate is much quicker.

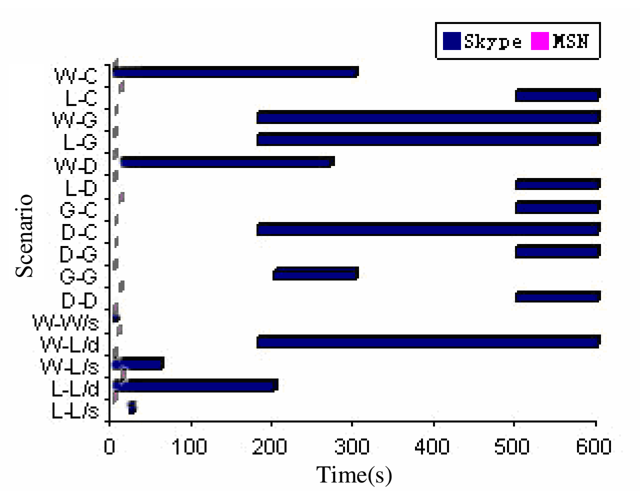

Figure 2. Presence comparison

Proceedings of the Advanced International Conference on Telecommunications and International Conference on Internet and Web Applications and Services (AICT/ICIW 2006)
0-7695-2522-9/06 $20.00 © 2006 IEEE

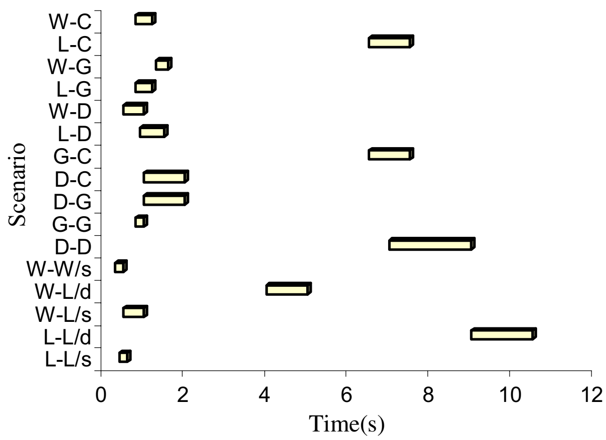

Figure 3. Presence of MSN Messenger

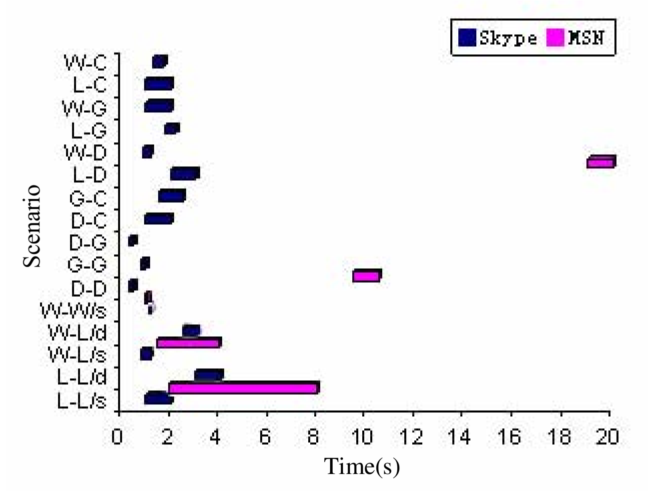

Figure 4. Connection Setup Delay comparison

3.2.2. Connection setup delay. Connection setup delay denotes connection speed. Its measure unit is *second* in this experiment. We measure the two kinds of software respectively in various scenarios. The results are given in Figure 4.

Generally, the Skype delay is shorter than MSN Messenger. Skype can work behind nearly any firewall and NAT, so it can work on almost every condition. But sometimes it is very hard to connect. Especially when Skype just begin to work, i.e. the user hasn't use Skype for a long time (actually only 1 day). After being used for a while, the status will be improved. When both sides work in different domains, the further the distance is, the longer the connection setup time costs.

For MSN Messenger, we can't set up connection in many scenarios, because it can't work through all firewalls. Our test shows that MSN Messenger can only work when both sides are in same domain or either isn't behind a NAT/firewall.

The connection setup delay is changeful due to the network condition and routing path.

3.2.3. M2E delay. M2E delay, i.e. mouth-to-ear delay, is one-way conversation delay. The voice is acceptable for most user applications when the delay is no more than 150ms [1]. Figure 5 shows the differences of M2E delay between Skype and MSN Messenger.

The result shows that the M2E delay of MSN Messenger is shorter than Skype. The delay of MSN Messenger is very short that usually we nearly don't notice it. The delay of Skype is obvious. But we can tolerate it in most scenarios. Both sides will have differences in M2E delay when either of us uses GPRS or CDMA. Because both GPRS and CDMA have different speeds between their uplink and downlink.

3.2.4. MOS. MOS is short for Mean Opinion Score. Generally it is obtained by asking a group of listeners to rank a speech sample on a scale of 1-5. In this experiment, due to restriction of headcount, we adopt multi-testing by one person to obtain MOS. The result is shown in Figure 6.

Generally, if MSN Messenger can setup connection, its voice quality is better than Skype. It has been proved that Skype works best in LAN, followed by DSL and CDMA. When it works with GPRS, the communication quality is very poor.

3.2.5. Handover Support. In wireless local area network, since the user can move around in the network coverage area, the seamless handover is a topic of importance. When a wireless user is moving and there are AP-to-AP (Access Point) handoffs in the network. Further delays are added into the WLAN as the user must associate with an AP, authorization must take place and the handoff must be completed [11, 12].

During the handoff between 802.11 APs, there will be short but noticeable loss of voice packets. On the positive side, VPN and call agent servers have timeouts on the order of tens of seconds. While some VoIP packets may be lost in a WLAN, connection should be maintained.

The comparison result is shown in Figure 7. From the result, we see that MSN Messenger is better than Skype in handover support.

For Skype, when AP switching occurs, there will be some blank periods of 3 to 7 seconds during conversation. When holding still, conversation will be continued after a while. But some words or phrases will be lost during the blank period due to the lost packets. When signal is too weak, connection will be lost.

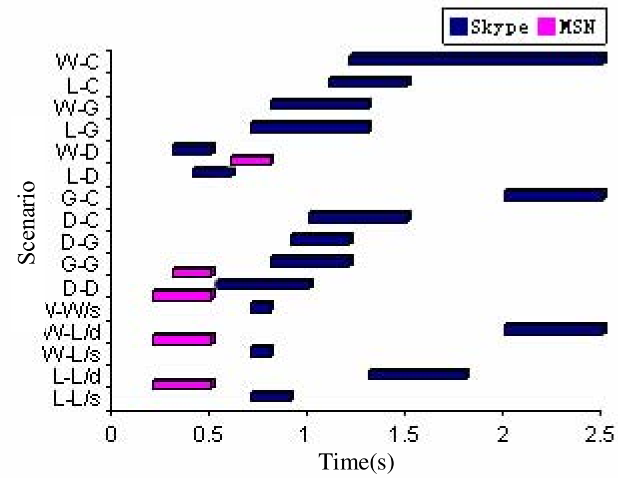

Figure 5. M2E delay comparison

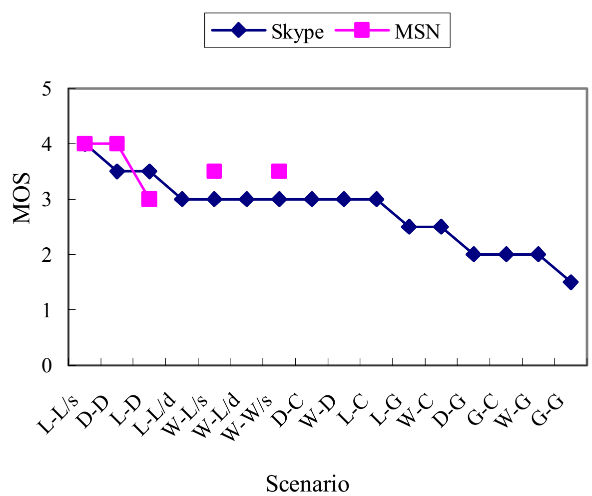

Figure 6. MOS comparison

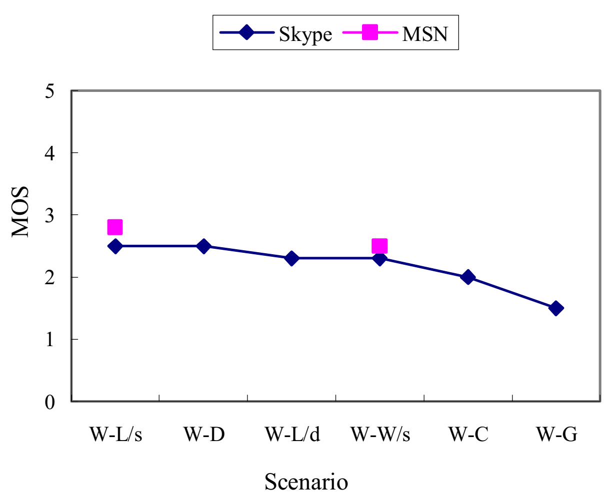

Figure 7. Handover support comparison

For MSN Messenger, when AP switching occurs during moving, the conversation will be slightly fitful, the blank periods during conversation is about 1 second. When holding still, the conversation will be continued. But some words or phrases will be lost.

4. Conclusion

Skype is the first VoIP software that introduces P2P technique. It becomes more and more popular for its several advantages such as high call completion rate, superior sound quality, configure-free, easy to use and secure communications. Skype is a P2P overlay network. It has a centralized database but decentralized traffic. It uses Global Index technology and GIPS iSAC Codec to deliver high quality telephony with lowest costs. Skype routes calls through multiple peers, to get around the firewall problem. Skype is able to intelligently route encrypted calls through the most effective path possible.

In order to compare the performance of VoIP software that uses P2P technique and other protocols. We make the performance comparison between Skype and MSN Messenger. Through the test result, we found that the voice quality of Skype is good but seems no better than MSN Messenger. Its delay is a bit longer than MSN Messenger but is still tolerable. When using P2P, its great benefit is that it can setup conversation wherever the two people are, because it can work almost seamlessly behind NATs and firewalls.

However, it still has many problems to be solved. We'll go on further study in the future.

Acknowledgement

This work was performed while Gao Lisha was a visiting student in Microsoft Research Asia. We're grateful to Zhu Wenwu and Zhang Qian for their helpful discussion and guidance. Special thanks to Ma Wenchao for his kind advice and discussion. This work is also supported in part by National Natural Science Foundation of China under Grants No.90412014, Jiangsu Provincial Key Laboratory of Network and Information Security under Grants No.BM2003201, and ChinaGrid Project of Ministry of Education of China under Grants No.CG2003-GA004-01.

References

[1] B. Goode, "Voice over Internet Protocol (VoIP)," *Proceeding of the IEEE*, Volume 90, Issue 9, Sept. 2002 Page(s):1495 - 1517

[2] Hong Liu and Mouchtaris, P., "Voice over IP signaling: H.323 and beyond", *IEEE Communications Magazine,* Volume 38, Issue 10, Oct. 2000 Page(s):142 – 148
[3] Schulzrinne, H. and Rosenberg, J., "The Session Initiation Protocol: Internet-centric signaling", *IEEE Communications Magazine*, Volume 38, Issue 10, Oct. 2000 Page(s):134 - 141
[4] Skype. http://www.skype.com/products/explained.html
[5] Skype Forum. http://forum.skype.com
[6] Salman A. Baset and Henning Schulzrinne, "An Analysis of the Skype Peer-to-Peer Internet Telephony Protocol", New York: *Columbia University Technical Report* CUCS-039-04, 2004
[7] Global IP Sound. http://www.globalipsound.com/partners/
[8] Lakaniemi A., Rosti J. and Raisanen V.I., "Subjective VoIP speech quality evaluation based on network measurements", *International Conference on Communications*, IEEE, no.1, June 2001, pp.748-752
[9] Matt Stoen, "New Perspectives on voice Quality", *Callisma whitepaper*, Oct 9, 2002
[10] Hassan M., Nayandoro A. and Atiquzzaman M., "Internet Telephony: Services, Technical Challenges, and Products," *IEEE Communications Magazine*, Volume 38, Issue 4, April 2000 Page(s):96 – 103
[11] Chyi-Nan Chen, Ching-Ta Lin, Juan-Min Zuan and Chung-Fann Liu, "The Study of Mobile Internet Telephony", *Proceedings. International Symposium on Multimedia Software Engineering*, IEEE, 2000. 11-13 Dec. 2000 Page(s):179 - 183
[12] Paul Struhsaker, "VoIP over WLAN in a Seamless Next-Generation Wireless Environment", *Texas Instruments whitepaper* SPLZ001 – June 2003.

Proceedings of the Advanced International Conference on Telecommunications and International Conference on Internet and Web Applications and Services (AICT/ICIW 2006)
0-7695-2522-9/06 $20.00 © 2006 **IEEE**