The Principles of Speech Transmission Realization in Skype

SOFIJA MIROTIĆ, IGOR RADUSINOVIĆ
Center for Telecommunication
University of Montenegro, Faculty of Electrical Engineering
Cetinjski put, 81000 Podgorica
MONTENEGRO

Abstract: - Skype is a peer-to-peer VoIP client that allows the use of voice and video communication, as well as sending of IM and files. Skype provides a possibility to overcome problems caused by the presence of NATs and firewalls. It uses an extremely reliable encryption system, so it is practically impossible to break into the connection or reveal the content of the message.

This article presents an analysis of peer-to-peer IP telephony using Skype application. Analys is is performed by monitoring Skype traffic in different network setups. The architecture of Skype is presented. Skype functions such as login, user searching, call establishment, media transfer and some technical features of Skype application are studied and described in this work. This report analyzes SkypeOut service which enables a Skype client to make a call towards a PSTN network subscriber.

Key-Words:- Skype, Peer-to-peer (P2P), IP telephony, Voice over IP (VoIP), Super node (SN)

1 Introduction

Skype [1] is a peer to peer VoIP client that meets the latest aspirations in the field of VoIP. It integrates several applications, i.e. it integrates voice mail, speech and video communication as well as exchange of IM (Instant Messages) and files. It appeared on the Internet in 2003 as the result of the joint work of two authors Nicholas Zennstrom and Janus Friis. Since then it has recorded very rapid developed reflected by the fact that each moment there are more than 3 million users on the Skype network. The reasons for its popularity can be found in the fact that Skype provides the best quality of speech communication in relation to other peer-to-peer VoIP clients. It also provides the possibility to overcome problems caused by the presence of NAT (Network Address Translation) and firewall. An important advantage of Skype is that it is supported by almost all operating systems (Windows, Linux, Mac OSX and Pocket PC).

Skype is private property and it is not an opened protocol which makes its analysis and the investigation of the functioning principle more difficult. The work [2] was the first one which presented the architecture of the Skype network and the basic principles of Skype protocol functioning. By applying the investigation methods described in [2] we reached new results. We determined changes in the Skype network architecture that appeared in the meantime i.e. that not only one Skype server is used for login but a number of them due to an increased number of users and reliability. We analyzed SkypeOut and determined that Skype has its own server for accounting. Also throughout the world there are servers that are not the property of Skype but have the role of gateway upon transition from Skype network (IP network) to the traditional PSTN network.

In this work we paid particular attention to the architecture of Skype network which we obtained by analyzing the process of connecting to Skype network, which is shown in chapter 2. Chapter 3 presents the functioning principle of Skype application i.e. the procedure of searching for a user, the establishing and the interrupting of a call. The manner of exchange of instant messages (IMs) and files is described in chapter 4. The principle of work of SkupeOut service is described in chapter 5. Chapter 6 compared Skype peer-to-peer VoIP client with the other currently topical VoIP clients. In the end conclusions are given and directions for future investigations defined.

2 Architecture and the basic characteristics of Skype

Skype is a peer-to-peer VoIP client that allows the use of voice and video communication, as well as sending of IM and files. With regard to the service it makes possible MSN [3], Yahoo messenger [4] and Free World Dailup (FWD) [5] are similar to it. However, they differ with regard to the protocol at the application level. The quoted applications base the transmission of the media content on the use of SIP (Session Initiation Protocol) [5] protocol, while the Skype client (SC) uses its own protocol.

In view of the fact that its predecessor was KazaA [6], the similarity in architecture is understandable.

Skype is an overlay peer-to-peer network, which means that peers form a logical network. This network is organized so that ordinary hosts are connected to super nodes which again are interconnected thus forming a network. The Skype network is actually made up of three basic elements:

- ordinary host,
- super node (SN),
- login server.

An ordinary host is every peer on which the Skype application is active. A super node can be every node that has a public IP address and that has a sufficient CPU, memory and network bandwidth. Its role is to connect the users and locate them on the network. Super nodes communicate among themselves. After connecting itself to the super node an ordinary host connects itself to the Skype's login server. This server has the role of ensuring authentication upon connecting to the network. Username and password are placed on the login server and are used at every logging on the Skype network.

We reached the results pertaining to the architecture of Skype network by monitoring and analyzing Skype peer-to-peer traffic. Taking an example by the practical part in work [2], we analyzed the Skype's traffic. We used Skype version 2.0.0.69 for the experiment, which we installed on two computers with Windows XP operating system. One PC is Intel Celeron 1.7 GHz, 120 MB RAM, and the second PC is Pentium IV, 504 MB RAM. For monitoring and analysis of the traffic we used sniffer analyzers Ethereal [8], Netpeeker [9] and LanExplorer [10]. In all three cases we obtained identical results. The content of signal messages at the application level cannot be revealed because of the use of encryption. Since a large number of hosts on the Internet are found behind NAT and firewall an experimental part pertaining to this topic was done. It is well known that NATs and firewalls are not an obstacle for Skype. It is assumed that Skype uses variant of the STUN protocol [9] in order to find out whether it is found behind NAT or firewall. However, there are no global NAT and firewall servers, but SC (Skype client) upon connecting on SN (Super node) obtains information related to the possible presence of NAT and firewall from SN. In cases when Skype client is found behind firewall which blocks UDP messages, all messages will be transmitted over TCP. However, in cases when firewall blocks also TCP messages at some ports, SC will try to realize a TCP connection with SN at first over port 80 (HTTP port), and if it does not succeed it will try on port 443 (HTTPS port). None of the firewalls blocks TCP messages on ports 80 and 443.

Taking example by the reference work [2] the following cases were observed:

- Both computers are with private IP addresses i.e. connected to the same LAN (100Mb/s Ethernet), and are found behind NAT.

- One computer is over dile-up connected by an ISDN line (one channel 64Kb/s) to the Internet and therefore has a public IP address, and the second one has a private IP address i.e. it is connected to LAN (100Mb/s Ethernet).

- Both computers are connected to the Internet over dile-up, one over a modem connection 56Kb/s, and the second one over ISDN and therefore they both have public IP addresses.

- One computer had a private IP address in network A which means it was found behind a port-restrected NAT and UDP-restricted firewall, while the second computer had a private IP address of network B. NAT and firewall used Linux as the operating system.

Skype uses an extremely reliable encryption system, so it is practically impossible to break into the connection or reveal the content of the message. Skype uses AES (Advanced Encryption Standard) [10]. In order to encrypt data in every Skype call or instant message, Skype uses the 256 bit encryption, which as a consequence has 1.1x 10^{77} of possible keys. Skype also uses the 1024 bit RAS to negotiate AES keys [1].

Precisely because of the encryption method it cannot be reliably determined which codec Skype uses. Skype automatically chooses the best manner of coding depending on the realized connection between two peers. Skype uses from 3-16 kb/s for transfer of speech depending on the available bandwidth and CPU the user has at disposal. In the period when it is ″free″, i.e. when there are no calls it uses from 0-0,5kb/s and this band is mainly used for transfer of information on the presence on the network. Since Skype uses the quoted bandwidth and at the same time offers a better quality of talk in relation to other peer-to-peer programs for VoIP it can be concluded that it uses some of the wide bandwidth codecs. It is assumed that Skype uses iLBC [11], iSAC [12] or some unknown manner of coding [2].

However, based on the exchanged TCP and UDP packages we reached certain conclusions on the network architecture. Upon the first logging on Skype network after installation of Skype application, SC establishes a TCP connection with the Skype server 212.72.49.131 on port 80 and sends GET for installation. After that it obtains the response HTTP/1.1 200 OK. Later with this server 212.72.49.131 on port 80 a TCP connection is established and GET is sent for checking the availability of the new version. This TCP connection is established at the end of every login on the Skype network no matter what ordinal number of login is in question. After exchange of TCP packages with Skype server 212.72.49.131 on port 80, SC sends UDP packages towards IP addresses of super nodes. The list of these IP addresses is taken over by SC immediately after installation, upon the logging on the network, during connection to Skype server. This list is automatically placed on the user's PC and SC builds and refreshes this list regularly. It can be found in the directory C:/ Documents and Settings/All Users/ Application Data/ Skype in the file Shared.xml. The list of super nodes actually includes the IP addresses and ports of these nodes. There are 200 IP addresses of super nodes on the list which were active at the moment of installation or at the moment of update

After sending UDP packages SC waits to receive the response from some of the super nodes to which it sent the UDP packages. SC establishes the TCP connection with the super node which sends a response first. TCP connection with SN remains during the whole of the period in which SC is active except in cases when the super node checks out from the Skype network. After establishing the TCP connection with SN, SC establishes the TCP connection with Skype login server. Most probably SC obtains from SN the information regarding the IP address of login server. We reached the conclusion that Skype uses a number of login servers and this result differs from the results in the reference work [2]. One group of servers is located in Danmark and their IP addresses are 195.215.8.141 and 195.215.8.142, and the second is the property of Skype and is located in Niderland and IP addresses are 212.72.49.141 and 212.72.49.142. Based on these conclusions the Skype network can be represented as in Figure 1.

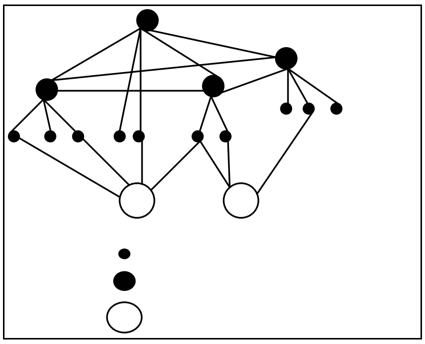

Fig. 1. Architcture of Skype network

Upon every login, no matter whether it was the first login after installation, or some of the subsequent login, the SC sends ICMP [13] messages. These messages are sent at the very beginning of logging, before UDP package is sent, as well as several times during the logging itself. ICMP messages are always sent to the same IP addresses, just like Figure 2 shows. These are ICMP messages of type 8 (Eho Request) and their role in the very login process is not clear.

As already stated, TCP connection between SC and SN lasts as long as SN is active. However, after the TCP connection is established with SN, SC continues to send the UDP packages towards other nodes and at the same time informs them on its presence on the network. Among the nodes to which the SC sends the UDP packages there are also those which are not on the list of super. When SN, to which the SC is coonected by TCP connection, ceases to be active its role is taken over by one of the nodes with which SC exchanged the UDP packages.

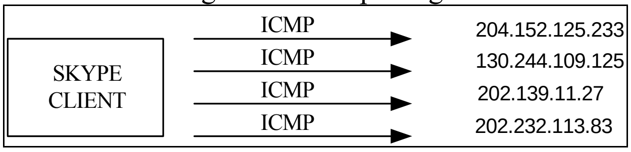

Fig. 2. ICMP messages sent by SC

These results are identical whether two SCs are started up on computers with public IP addresses or on computers which have private IP addrsses. Figure 3 shows the login process in the quoted cases.

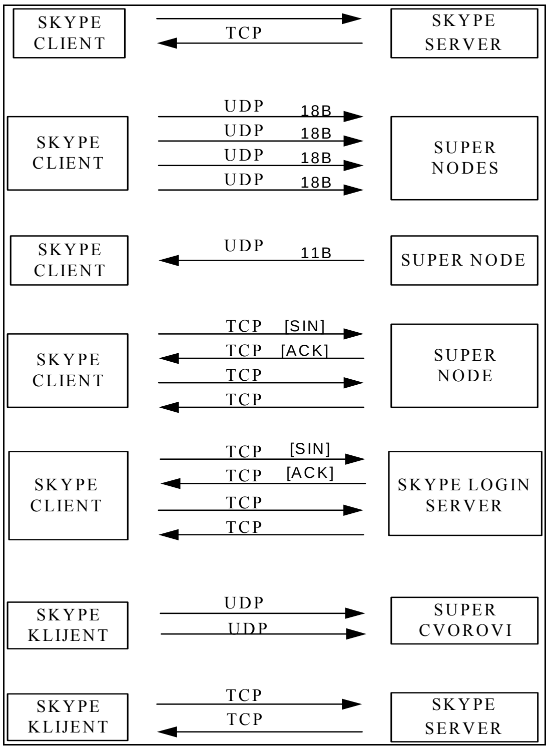

Fig. 3. Login proces of SC to the Skype network when none of the SCs are behind the firewall

Difference in login process to the Skype network was observed only in the case when one SC is in LAN A behind the NAT and the firewall which blocks the UDP packages. In that case upon the first login after installation of Skype SC first established the TCP connection with the Skype server for the purposes of installation, then it sent a number of UDP packages towards different IP addresses for which we asumed to be super nodes. As it cannot receive UDP packages from super nodes, SC waits for some time and after that sends the TCP package towards the first IP address from the range. After obtaining a confirmation of the receipt of the TCP package, the TCP connection is established with that super node. After that follows the establihsment of the TCP connection with the Skype login server.

3 User search, call establishment and teardown

In order to allow successful search in the distributed peer-to-peer network Skype uses its own technology, the so called Global index technology or the third generation of P2P technology (3GP2P) [14]. The Global Index technology is multy-tiered network in which the super nodes communicate among themselves and with the minimum delays have at their disposal information on the presence of all users and on the availability of resources on the network. The authors of Skype claim that every user who was active on the network in the past 72 hours can be found.

During a process of user search the TCP messages are exchanged between SC and SN, after which SC sends UDP packages towards several IP addresses. This procedure is repeated until the user is found or until information that the user does not exist is received. First UDP packages are sent towards a certain number of IP addresses (most frequently 6), if the user is not found TCP packages are again exchanged between SC and SN, then SC sends the UDP packages towards new IP addresses. Usually the number of IP addresses is increased and this increase in the number of IP addresses to which UDP packages are sent continues until the final response to the set search is obtained. It is assumed that in TCP messages the SN gives addresses of possible SNs to which the searched user could be connected. After the search is completed, TCP messages are exchanged between SC and SN. When the searched user is found he is added to the buddy list after that he can be forwarded a call if he is currently on the network.

Regarding the buddy list, at the beginning it was found placed only on the machine from which Skype application is initiated, which required the user to recreate the buddy list in case he uses a number of machines for login on the Skype network. However, already after Skype version 1.2 (March 2005), the buddy list was placed on the Skype server and it can be accessed from every computer.

In case when both SC clients are activated on machines connected to the same LAN network and are found behind the NAT without the presence of firewall, upon establishment of a call the signal messages are exchanged directly between two peers over TCP, while the media content is transmitted between them over UDP, which is shown in Figure 4. Also during talks in a particular interval TCP messages are exchanged between two peers, the aim of these messages being to determine the status of the speech connection.

When both SCs have public IP addresses, signal messages being exchanged upon establishment of the call are sent over TCP. However, these TCP messages are exchanged with the super nodes which are found between two peers, and not directly between the peers themselves, while the media content is exchanged by UDP directly between peers, as Figure 5 shows.

When one SC is connected to LAN behind NAT, without the presence of a firewal, and the second SC has a public IP address, signal messsages for establishment of a connection are exchanged over TCP over intermediary nodes. However, the media content is sometimes exchanged over UDP directly between two peers as in Figure 5, and sometimes also over UDP, but over intermediary nodes over which TCP messages have already been exchanged. This case is described in Figure 6. The manner of media content transfer probably depends on the degree to which the Skype network is burdened.

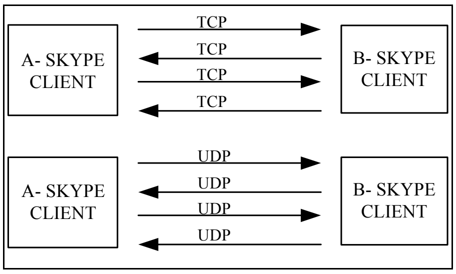

Fig. 4. Process of call establishment between two Skype clients connected to the same LAN

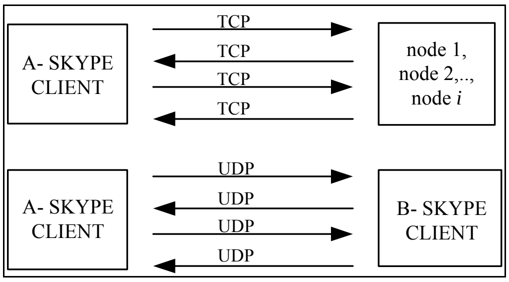

Fig. 5. The proces of call establishment between two SCs when they both have public IP addresses

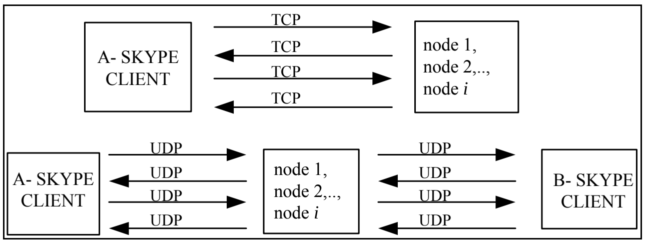

Fig. 6. Call establishment between two SCs when one is in LAN, and the other has a public IP address

Even when there is no speech, UDP messages are exchanged, which means that there is no silence suppression. During the talks TCP messages between SCs and SNs are also exchanged in the interval of 60 seconds. These signal messages are used to determine whether a peer is still on the network i.e. whether it is active.

When one SC is found in LAN A behind NAT and firewal, and the second SC is found in LAN B behind NAT without the presence of firewall, signal messages for establishment of connection and media content are transferred over intermediary super nodes by TCP connection as Figure 7 shows.

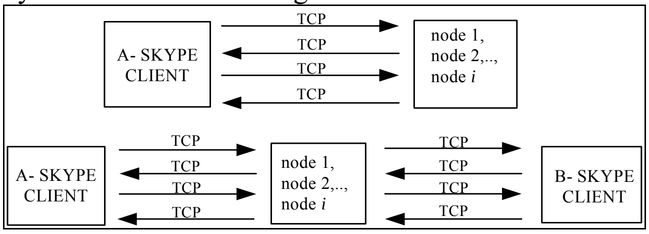

Fig. 7. Call establishment between two SCs where one is in LAN A and the other SC is in LAN B

Skype provides the possibility of up to five user conference connection though there is an agreement between the Skype company and Intel to make possible a ten user conference connection for the future Intel generations of laptops and PC computers. Skype also makes possible diverting of calls as well as putting of a call on hold. In distinction from other peer-to-peer VoIP cients, such as YahooMessenger and MSN, Skype provides the possibility of registering the same user to Skype network with several computers with the same user name and *password*. In that case the call directed to him can be received from any PC. When the call is taken over from one computer signalisation of the incoming calls on other computers ceases.

4 Sending of instant messages and files

When Instant Messages (IM) are sent, TCP is used for transfer of messages, this we conclude based on the size of TCP package. However, for transfer of files UDP is used just like in the case when media are transferred. This is naturally valid in case when both SC clients were activated on machines with public IP addresses or one or both SC clients are found behind NAT without the presence of a firewall.

Complete exchange of IM between two SCs is placed on every computer in the directory C:/ Documents and Settings/ All Users/ Application Data/ Skype. If we are sending IM to a user who is logged on several computers simultaneously, he will receive messages on all the computers.

5 Establishing calls between Skype and the user of PSTN network (skypeout)

Apart from free of charge services, Skype also offers to its users some services that are being charged. This is the SkypeOut service which enables a Skype client to make a call towards a PSTN network subscriber and was introduced in June 2004. In March 2005 SkypeIn was introduced, which enables calling from PSTN network of a Skype user, who receives the call on his PC. When the Skype user activates this service, he is allocted a subscriber's number to which PSTN users can call him. Since March 2005 there is the possibility of using voice mail.

Within the practical part of this work, special attention was paid to SkypeOut service. Since Skype is a peer-to-peer VoIP protocol we analyzed the call made from Skype client towards PSTN network. The analyzed case was the one when SC on LAN is behind NAT, but without the presence of a firewall, as well as the case when SC has a public IP address. The calls were made towards PSTN networks of countries in all continents.

When Skype application is already active and the calling of the subscriber's number on PSTN network begins, at the very beginning already, apart from the already existing TCP connection with SN, TCP connection with Skype server is also established with the aim of adequate accounting of the call. These are most frequently servers whose IP addresses are 212.72.49.155 (which belong to Skype-NL whose address range is from 212.72.49.128 to 212.72.49.159) and 195.215.8.140 (which is located in Danmark and belongs to the company whose address range is from 195.215.8.0 to 195.215.8.255). Signal messages for the establishment of the connection are exchanged by TCP with some of these servers. The media content is exchanged directly between the host on which the Skype application is active and the gateway by UDP. The exchange of TCP and UDP messages is shown in Figure 8.

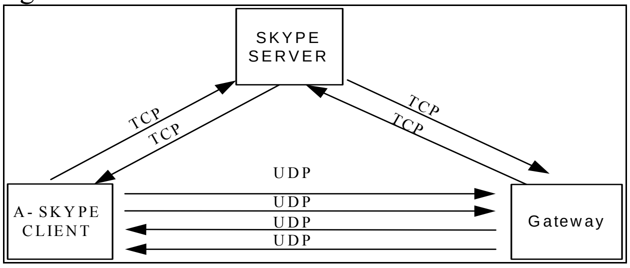

Fig. 8. Establishment of a speech connection from SC client to PSTN network

When making calls from SC to PSTN network the media content is sent towards different gateways depending on the location of the subscriber. Thus calls towards the PSTN network in Europe ended at *gateways* whose IP addresses are from the range (213.244.170.64 -213.244.170.127) and (80.92.84.0 - 80.92.85.0). Calls towards the PSTN networks in other parts of the world ended at IP addresses: (216.168.164.116),(216.168.164.118), (198.173.5.23), (67.72.71.50), (4.79.96.176), (61.66.230.247). For all that there is no division according to continents or areas, but for example the same gateway is used for some countries of Asia and some countries of South America.

6 Comparison of Skype with competitive P2P VoIP applications

Within this work we analyzed also other applications which enable speech and video communication, exchange of IM, sending of files, such as Free World Dialup, MSN, Yahoo messenger. FWD for VoIP signallization uses SIP protocol. This VoIP client has the possibility of audio and video communication, sending of instant messages (IM) and conference call. In distinction from Skype it does not include the possibility of sending files. The possibility of dialing PSTN number from a PC, as well as of receiving calls from PSTN network on PC in FWD is still in the development phase. NAT and firewall are not an obstacle for FWD, it uses STUN. FWD has STUN server at its disposal (stun.fwdnet.net port 3478).

MSN client for VoIP signallization uses its own MSNMS/TCP and SIP/UDP protocols, and for transfer of media it uses RTP/UDP protocols. It provides the possibility of speech and video communication, exchange of IM and exchange of files. In distinction from Skype it does not include the possibility of conference call and there is no possibility to use voice mail.

Yahoo messenger is a P2P VoIP application which enables the use of speech and video communication, exchange of IM and files, it also enables conference call, as well as voice mail in case when the user if offline. In distinction from Skype the voice mail service is free of charge. Yahoo messenger has also the possibility of using the services Call In and Call Out which are charged to users. Call Out is the service which makes it possible to the users to make calls from the PC to PSTN network, while the Call In service makes it possible to the user to receive the call from PSTN network to PC.

For VoIP signallization Yahoo messenger uses SIP protocol. For transer of media content it uses the RTP/UDP protocol. Upon login, sending of IM and sending of files this client uses its own protocol YMSG at the level of the application, and TCP at the level of transport. It uses STUN protocol for overcoming the problems caused by the presence of NAT.

7 Conclusion

Skype is a peer-to-peer VoIP client which in large exceeds the competitive clients in terms of its possibilities and quality of communication it provides to the users. This is at the same time the reason for which it has developed so rapidly and why it is ever more present in peer-to-peer VoIP communication. Due to the fact that this is not an open protocol, it is the subject of large interest of experts for VoIP technology and protocols.

The aim of this work was to present the architecture of Skype network. Through analysis of Skype traffic we reached the conclusion that Skype uses a number of login servers, probably because of an increased number of users and redundance. This work presents the manner in which calls from Skype clients to PSTN network are realized. The existance of Skype server to which the Skype client is registered when the call towards the PSTN network is established and his role in tariffing is determined.

The plans for future investigations are deeper analysis and investigation of Skype protocol.

References:

[1] Skype http://www.skype.com

[2] Salman A. Baset, Henning Schulzrinne *An Analysis of the Skype Peer-to-Peer Internet Telephony Protocol,* September 2004.

[3] MSN http://www.msn.com

[4] Yahoomessenger http://messenger.yahoo.com

[5] FWD http://www.freeworlddialup.com

[6] J.Rosenberg, H. Schulzrinne, G. Camarillo, A. R. Johnston, J.Peterson, R. Sparks, M. Handley, and E.Schooler. *SIP:Session Initiation Protocol*. RFC 3261, IETF, June 2002

[7] KaZaA http://www.kazaa.com

[8] Ethereal http://www.ethereal.com

[9] Netpeeker http://www.netpeeket.com

[10]LanExplorer http://www.sunrisetelecom.com/lansoftware/lanexplorer.shtml

[11] J. Rosenberg, J. Weinberger, C. Huitema, and R. Mahy. STUN: *Simple Traversal of User Datagram Protocol* (UDP) *Through Network Address Translators* (NATs). RFC 3489, IETF, March 2003

[12] AES Advanced Encription Standard http://csrc.nist.gov/publications/fips/fips197/fips-197.pdf

[13] S. Andersen, A. Duric, H. Astrom, R. Hagen, W. Kleijn, J. Linden *Global IP Sound, Internet Low Bit Rate Codec (iLBC)*, December 2004

[14] iSAC code www.globalipsound.com/datasheets/iSAC.pdf

[15] J. Postel, Internet Control Message Protocol, September 1981

[16] Global Index (GI) http://www.skype.com/skype_p2pexplained.html