The Secrets of Skype Login

Po Qi
School of Computer Science and Technology Harbin University of Science and Technology,
Harbin, 150080, China
poqi@tsinghua.edu.cn

Cuilan Du, Yan Ren
CNCERT/CC China
Beijing, 100029, China
ducuilan@cert.org.cn,
renyan@cert.org.cn

Yibo Xue*
Tsinghua National Laboratory for Information
Science and Technology
Beijing, 100084, China
yiboxue@tsinghua.edu.cn

***Abstract*—Skype is beyond any doubt the most popular VoIP application on the Internet. Its amazing success has drawn the great attention of telecom operators and the research communities, both are interested in knowing its internal mechanism, characterizing its traffic, and understanding its users' behavior. However, due to its proprietary, these problems have not been completely solved, which likes a black box that hides a lot of secrets. Skype login is the first step of the whole communication process and very critical for guaranteeing the Quality of Service. This paper aims to reveal the secrets of Skype login. Based on a lot of experiments and observations, we deeply analyze the Skype login, and completely make clear the various functions, traffic characteristics and login process of Skype. Uncovering these secrets is very useful for protocol identification, traffic classification, network management and optimization.**

Keywords: Skype Login; Login Server; Skype Client; Protocol Identification, Traffic Classification

I INTRODUCTION

In the last few years, we witnessed the Voice over Internet Protocol (VoIP) gaining a tremendous popularity, which be testified by the increasing number of operators that are offering VoIP-based phone services. Skype is a typical peer-to-peer application as a successful commercial VoIP system [1], and becomes more and more popular on the Internet because of easy-to-use and high quality [2]. As so far, its user number has reached up 225 million. It has been grown to be a strong competitor to the traditional Public Switched Telephone Network (PSTN). Therefore, both in terms of the number of users and the influence in the field of VoIP, Skype has become one of the most popular application on the Internet today.

Being the most popular and successful VoIP application, Skype is attracting the great attention of the research communities and the telecom operators. Many interesting questions are involved in its internal mechanism [2][3][4], traffic features [5] and users' behaviors. However, it is very hard to get these answers, because Skype is a proprietary protocol, and adopts cryptography and obfuscation techniques for anti-reverse engineering [6]. In addition, Skype implements a number of techniques to circumvent NAT and firewall limitations, which add further complexity to an already blurred picture [2]. Therefore, Skype hides a number of secrets and poses significant challenges for telecom operators and the security researchers. Up to now, these secrets have not been completely revealed.

If Skype client (SC) is able to accomplish the normal communication, it should login successfully. Consequently, Skype can be divided into the login stage and communication stage that consists of text chat, file transfer, audio, video, and so on. Skype login is the most critical step for Skype communication. Because all the Skype traffic is encrypted, the secrets including the login process, the necessary function and the traffic Characteristics are hidden [7], and are not still fully revealed.

In this paper, we successfully reveal the secrets of Skype login by analyzing a mass of the Skype network traffic.

The secrets of Skype login are as the following:

- *Login Process*: We completely reveal the secrets of the login process and describe the various functions performed by each step in details.
- *Traffic Characteristics:* During the Skype login process, SC tries to establish a connection with various types of servers to complete some function, which creates some fixed traffic that contain some features. So we revealed the flow characteristics. For example, port (fixed port, the range of port), packet information (packet size, the number of packets), protocol (Transmission Control Protocol, User Datagram Procotol), and so on.
- *Server Resources:* We discover some important servers provided by Microsoft and Skype company, and expound the characteristics of these servers. For example, Ethernet IP address, country, city, organization, company, operating system, device type, open port, service and software version, and so on.

Revealing the secrets of Skype login is like opening a black box, these uncovered secrets can provide a theoretical basis of traffic classification and protocol identification, and validated the idea of protocol behavior chain [8].

The rest of the paper is organized as follows, Section II describes the related works about the study of Skype login. Section III describes the secrets of Skype login that consist of login process and the various functions. Section IV summarizes our findings and future work.

II. RELATED WORK

In the past few years, some researchers have been promoted a lot of work in analyzing the protocol of Skype [2][3][4]. They have achieved some research results,

*Corresponding author. Tel: +86 010 62772393. E-mail: yiboxue@tsinghua.edu.cn

including key Skype functions such as login, NAT and firewall traversal, encryption, codecs, call establishment, and totalizing the number of super nodes in the global, and so on. Because of the requirements of the network optimal management, some researches focus on the identification of the Skype network traffic [9][10][11]. But there is not an effective method to be able to completely identify Skype traffics now.

Although some scholars have done some researches on Skype login, they already revealed some features of Skype login, such as login process, login server, traffic characteristics and login process time [2][3][4][11], but they did not reveal all and did not accurately describe the important server resources. In 2012, Microsoft acquired the Skype Company and integrated Microsoft Service Network (MSN) into Skype. Skype is now able to share friend list with Facebook and constantly renews new version. This results in new changes, so the secrets that consist of flow characteristics, internal mechanisms and login process of Skype login has certainly changed.

This paper will reveal all the secret of Skype login, it will be a theoretical foundation for further research on Skype.

III. THE SECRETS OF SKYPE LOGIN

Skype login is the most complex operation in the Skype. Because all the Skype traffic information is encrypted, many secrets are not exposed in the Skype login process, which consist of the login process, function and traffic characteristics, and so on.

After a lot of experiments were performed for the Skype version 5.6.0.106 and 6.3.0.197, we collected a large number of Skype traffic. Based on a lot of deep analysis and observations, we reveal the secrets of Skype login. In this section, we elaborate on the secrets of Skype login.

A. Login Process

According to the time sequence, the login process exactly tells us every step of the operation in the Skype login process. It has some fixed functions. For example, SC needs to authenticate its user name and password with the login server, discover online Skype nodes with public IP addresses, determine the type of NAT and firewall it is behind, advertise its presence to the others peers and its buddies, check the availability of the latest Skype version and establish a communication channel with the user feedback information server and Facebook server. The Fig.1 completely describes the login process.

As illustrated in Fig.1, each server group has a set of servers and provides the same function or service. SC needs to access to several groups according to the fixed sequence. This sequence is as follows.

1. SC checks user name and password with Login Server.
2. SC establishes a connection with Skype server and makes communication with four super nodes that are randomly selected from the local file, which ensure that SC certainly connects to a super node.
3. SC sets up communication with MSN Echo Server to determine the type of NAT and firewall it is behind.

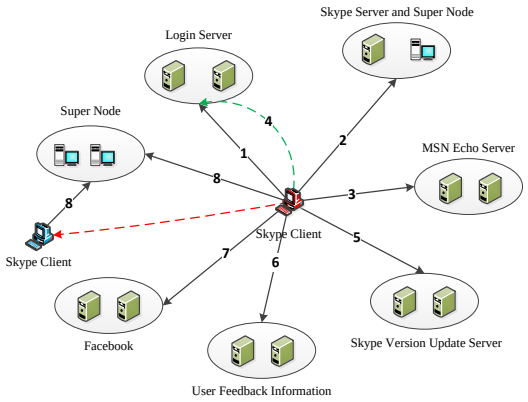

Figure 1. The Login Process of Skype Login

4. SC advertises its presence to the others peers and its buddies. If user's buddies are not stored in the system local file, SC should obtain the friend list from the login server and advertise its presence.
5. SC gets the latest Skype version.
6. SC keeps active communication with the user feedback information server to collect user suggestions.
7. SC tries to connect Facebook server in order to share the friends on Facebook, this can achieve a multi-account applications on different platforms.
8. If SC wants to communicate another SC, it first connects the super node and then directly connects to another SC.

B. The Various Functions

There are various functions during Skype login, each group servers finish the same function. We elaborate on them in the following, then a number of secrets are uncovered one by one.

1) Verifying the User Name and Password

Login servers are the only central component in the Skype peer-to-peer network. They store Skype user names and passwords, and ensure that Skype user names are unique according to the Skype name space. SC must first authenticate itself with the login server for a successful login.

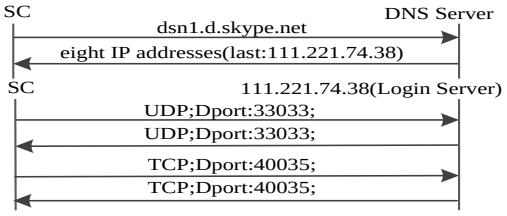

Figure 2. The Process of Verifing the User Name and Password

Fig.2 shows the process between SC and login server. During the login process, SC firstly selects one of eighteen domains (dsn0.d.skype.ent-dsn17.d.skype.net), then sends the selected domain name request to the DNS server. The

DNS server will return eight IPs. Secondly, SC selects the last IP as the login server. Lastly, SC establishes communication with the selected login server using the UDP protocol on port 33033 and maintains the communication with login server using the TCP protocol whose port is one of 47 ports (40041-40047).

We verify that there are 225 login servers distributed in the world, as shown in Table I.

TABLE I. DETAILED DESCRIPTION ON SKYPE LOGIN SERVER RESOURCES

C-Class	IP	Count	Organization	Country
111.221.74	111.221.74.12-34,37-28	25	Microsoft Hosting	Singapore
111.221.77	111.221.77.140-162,165-166	25	Microsoft Hosting	Singapore
157.55.130	157.55.130.140-162,165-166	25	Microsoft Corp	United States
157.55.235	157.55.235.140-162,165-166	25	Microsoft Corp	United States
157.55.56	157.55.56.140-162,165-166	25	Microsoft Corp	United States
157.56.52	157.56.52.12-34, 37-38	25	Microsoft Corp	United States
213.199.179	213.199.179.140-162,165-166	25	Microsoft Data Center	Ireland
64.4.23	64.4.23.140-162,165-166	25	Microsoft Hotmail	United States
65.55.223	65.55.223.12-34,37-38	25	Microsoft Hosting	United States

As illustrated in TABLE I, these 225 login servers belong to Microsoft Company, which distributed in the global. They are allocated to 9 C-Classes, Each C-Class has 25 IPs. They have same feature as shown in Table II.

TABLE II. THE FEATURE OF SKYPE LOGIN SERVER

OS	Device Type	Open Port	Service	Service Version
Linux 2.6.23	PBX | WAP	80, 443, 33033, 40001-40047,	Skype	Skype2

TABLE II shows the feature of 225 Skype login servers, they have the same OS, device type, open port, service and service version.

The complete communication process between SC and Login Server is illustrated in Fig.3. SC selects one of eighteen domains (dsn0.d.skype.ent-dsn17.d.skype.net) and sends the selected domain name request to the DNS server. The DNS server returns eight IPs. Then, SC firstly sends a UDP packet to each of the eight IPs on port 33033. If there is no response, SC then sends three TCP packets with 62 bytes length to each of the eight IPs on port 33033. If there is no response yet, SC tries to establish a TCP connection using three packets with 62 bytes length to each of eight IPs on port 80 and port 443. If any one of the above actions is successful, Login is passed. Otherwise, the connection is failed, it will repeat the whole process after 15 seconds.

2) Connecting the Super Node

SC must connect to a super node for communication, so Skype usually writes the default fixed super nodes provided by the Skype company into client code. After SC finishes checking user names and password, it makes a connection to one Skype server using the TCP protocol on port 33033.

There is a hidden file under the folder "Protected Storage", its path is /HKEY_CURRENT_USER/Software /Skype/Protected Storage at the current path of registry table. The file contains a list of super nodes, it is refreshed regularly. During this stage, SC randomly selects four super nodes from the file to make communication using UDP protocol.

Both Skype server and four super nodes are able to guarantee an available and stable connection.

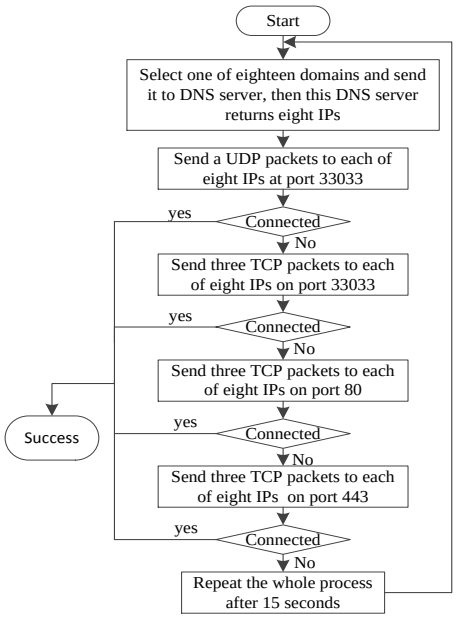

Figure 3. The Process between SC and Skype Login Sever

3) Determining the Type of NAT and Firewall

We conjecture that SC uses a variation of the STUN and TURN protocols to determine the type of NAT and firewall it is behind. However, the MSN Echo Server realizes the technique to determine the type of NAT or firewall in the login process, which guarantee a stable peer-to-peer communication.

InFig.4, SC first sends a domain name request whose name is s.gateway.messenger.live.com to the DNS server, and the DNS server returns one IP (such as 65.5.121.231). Then SC makes a communication with the MSN Echo Server using TCP protocol on port 443.

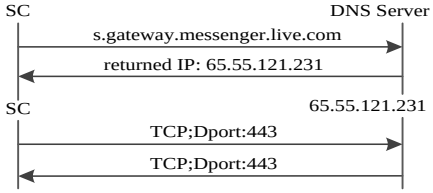

Figure 4. The process Between SC and MSN Echo Server

4) Advertising Its Presence

Login servers store not only the user names and passwords but also buddy list. So SC advertises its presence to other peers and its buddies by the login server. If SC did

not have the buddy list in the local file, SC should get it from the login server in the login process.

Skype uses its Global Index (GI) technology to search for a user. Skype claims that the search is distributed and guaranteed for finding a user if it exists and has logged in during the last 72 hours. These 225 login servers use the GI technology.

5) Skype Version Updating

During the login process, SC sends an HTTP 1.1 GET request to the Skype Version Update Server (ui.skype.com) to check whether a new version is available. The first line of this request contains the string 'get the latest version'. During the first HTTP request, there are only text-based messages sent by Skype.

Fig.5 shows the process between SC and Skype version update server.

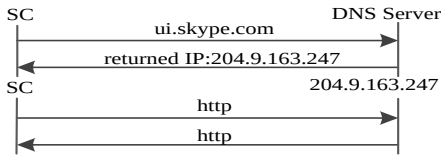

Figure 5. The Process between SC and Skype Version Update Server

6) User Feedback Information

User Feedback information Server aims to collect users' feelings and proposal which includes users' experience, instant messages, phone, video, new function and using frequents on Skype every week, this helps to promote the Skype the users' experiences.

Fig.6 shows the process between SC and User Feedback Information Server. SC firstly randomly selects one of fourteen domains (166.0.10.5.rst0.r.skype.net-166.0.10.5.rst 13.r.skype.net) and sends its request to the DNS server. Then the DNS server returns eight IPs, SC establishes a communication with the User Feedback Information Server using TCP protocol on port 12350.

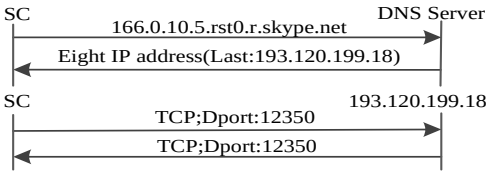

Figure 6. The Process Between SC and User Feedback Information Server

7) Connecting the Facebook

During login, SC tries to connect the Facebook servers and share the buddy lists of the Facebook. At first, SC sends a domain name request of which name is connect.facebook. com to the DNS server. The DNS server returns IP addresses, then SC makes communication with Facebook using TCP protocol.

8) Connecting Another SC

If SC wants to communicate with others SC, it firstly connects to a super node, then directly connects to the SC. Because any node with a public IP address, sufficient CPU, memory, and network bandwidth can become to be a candidate of super node, super node is randomly allocated by network conditions.

Hereto, we reveal all the secrets of Skype Login, it is very useful for protocol identification, traffic classification, network management and optimization.

IV. CONCLUSION AND FUTURE WORK

We have done a lot of experiments and collected huge Skype traffic. Based on a deep analysis and observations, we revealed the secrets of Skype login, uncovered the detailed information and process of the login process, internal mechanism, traffic characteristics (e g. domain, port, protocol, server IP address and the feature of server device and so on) and server resources, this can provide a theoretical basis for traffic classification and protocol identification.

In the future, we will analyze Skype communication process, and try to reveal the secrets of the Skype communication process. This will help to completely understand the internal mechanism and users' behavior of Skype, after that we will design a Skype Network Traffic Detection System (SNTDS) in order to completely identify the Skype traffic.

ACKNOWLEDGMENT

This work was supported by the National Key Technology R&D Program of Chnia under Grant No.2013BAH46B04.

REFERENCES

[1] Skype – the whole world can talk for free, information on http://www.skype.com.

[2] S. A. Baset and H. Schulzrinne, "An analysis of the Skype peer-to-peer internet telephony protocol," In IEEE infocom Vol. 6, pp. 23-29, April 2006.

[3] Guha, Saikat and N. Daswani, "An experimental study of the Skype peer-to-peer VoIP system," Cornell University, 2005.

[4] Biondi, Philippe and F. Desclaux, "Silver needle in the Skype." Black Hat Europe, 2006, pp. 25-47.

[5] Bonfiglio, Dario, et al. "Detailed analysis of Skype traffic," Multimedia, IEEE Transactions on 11.1, 2009, pp.117-127.

[6] J. Liang, R. Kumar and Kw. Ross, Understanding KaZaA. Information on http://cis.poly.edu/~ross/papers/Understanding KaZaA.pdf, 2004.

[7] Bonfiglio, Dario, et al. "Revealing Skype traffic: when randomness plays with you," ACM SIGCOMM Computer Communication Review. Vol. 37. No. 4. ACM, 2007.

[8] Luoshi Zhang, Dawei Wang and Yibo Xue, "BCBPI: A Novel Behavior Chain Based Protocol Identification Method," Cloud Computing and Information Security, International Conference, 2012.

[9] Korczynski, Maciej, and Andrzej Duda. "Classifying service flows in the encrypted Skype traffic," Communications (ICC), 2012 IEEE International Conference on. IEEE, 2012.

[10] Meng Zhang, et al. "Encrypted Traffic Classification Based on an Improved Clustering Algorithm." Trustworthy Computing and Services. Springer Berlin Heidelberg, 2013. pp.124-131.

[11] C. Macharaschwili and L. Skidmore, "A Skype-Buddy Model for Blended Learning," Journal of Interactive Learning Research, Vol. 24 (2),2013,pp.167-19.