A Generic Technique for Voice over Internet Protocol (VoIP) Traffic Detection

Fauzia Idrees    Uzma Aslam Khan
Military College of Signals, NUST
Rawalpindi, Pakistan

Summary
Skype, Google Talk, Yahoo voice etc. are all applications that enable the use of the Internet for voice conversations. They offer cost effectiveness and are easy to use, and due to these reasons many new VoIP applications are coming into existence. However, all forms of communications need to be monitored for security purposes to ensure their correct usage. With the development of more and more VoIP applications, monitoring and detection of these applications is becoming a more difficult task. Most detection techniques are based on standard protocol and IP address identification. Thus, application detection and monitoring techniques are developed after an application has been in use for some time, resulting in obvious security implications. This paper presents generic techniques for the detection of traffic generated by all VoIP protocols, both currently in existence and any future VoIP protocols that may be used. The method proposed is based on analysis carried out on different VoIP applications currently in existence.

Key words:
Voice over IP (VoIP), Skype, Peer-to-Peer (P2P), Internet Telephony, Voice packet characteristics

1. Introduction
The world is becoming increasingly IP-centric, with a large number of devices getting networked every day. Voice over Internet Protocol -VoIP is one of the fastest growing Internet applications today. Voice over IP -VoIP - is a set of technologies that enable voice calls to be carried over the Internet (or other networks designed for data), rather than the traditional telephone landline system—the Public Switched Telephone Network (PSTN). Voice over IP uses the Internet Protocol (IP) to transmit voice as packets over an IP network. Using VOIP protocols, voice communications can be achieved on any IP network regardless, it is Internet, Intranets or Local Area Networks (LAN). The potential of free or very low cost- phone calls is the driving force behind the adoption of this technology, but in the long run, VoIP is more significant than just free phone calls, it represents a major change in telecommunications. The fact that VoIP transmits voice as digitized packets over the Internet means that it has the potential to converge with other digital technologies, which in turn will result in new services and applications becoming available.

VoIP is an advancing area of research. There are many different and generally incompatible techniques for sending voice over the Internet. The International Telecommunications Union standard H.323 provides for voice and video teleconferencing; the Internet Engineering Task Force adopted an incompatible system called Session Initiation Protocol (SIP). Cisco developed a proprietary system called the Skinny Client Control Protocol (SCCP). This variety of available protocols has led to several different implementation architectures. Most implementations use the centralized server client architecture, but recent years have also seen developments in the decentralized peer-to-peer networks.

Offering a cost effective solution without a compromise to the quality is attracting both home users as well as businesses, which are dependent on long distance communications. A recent survey carried out predicted that VoIP will account for approximately 75% of world voice services by 2008.

However, the adoption of VoIP is not without its complications. Law enforcement agencies often need to conduct lawful electronic surveillance in order to combat crime and terrorism. The telephone service provider is required to provide the authorized law enforcement agencies with contents of telephone calls conducted by each user designated for surveillance. Carriers want to identify the type of traffic their networks are carrying, especially VoIP calls. The emphasis on VoIP is because it uses up the carriers’ largest traditional source of revenue, circuit switched services. Even if they offer VoIP services themselves, they face an obvious dilemma. At the very best, they receive less revenue from their largest and most
VoIP detection is an emerging field. Different techniques have been developed for identifying the VoIP communications by different researchers and commercial organizations. However, most of these solutions are either specific to some protocol or based on predefined ports.

Port based analysis is the most basic and straightforward method to detect P2P users in network traffic. It is based on the simple concept that many P2P applications have default ports on which they function. When these applications are run, they use these pre-defined ports to communicate with outside. However since VoIP services use no specific port rather voice packets communication is accomplished through dynamic ports to avoid the detection. Moreover, some applications also masquerade their functional ports as well-known application ports such as port 80. Due to these points there is a greater possibility of getting false positives while detecting applications through port based analysis. All these issues make port based analysis less effective for VoIP detection.

In protocol based approach, a software or hardware tool is used to monitor the traffic passing through the network and investigates the data payload of the packets according to some previously defined application signatures specific to the underlying protocol of application. Many of today's commercial and open source P2P application detection solutions are based on this approach. They each accomplish the detection by looking for the expression matches on the application layer data, in order to determine whether a special P2P application is being used.

Because protocol analysis focuses on the packet payload and raises alerts only on a definite match. As VoIP applications are emerging incessantly, and therefore new signatures also keep adding up. Static signature based matching therefore requires continuous and timely updating to cater for the new protocols. VoIP developers are also inclined to circumvent detection by encrypting the traffic, making protocol analysis much more difficult.

As stated by Curtis [5] “Unlike other traffic types VoIP cannot be simply identified by IP fields or by port usage. Also, because measurements have been taken on real operating networks, security concerns have meant that only header information is available not packet contents. Thus, identifying VoIP traffic is a non-trivial task”.

Identification of VoIP traffic was carried out by J.P. Curtis, J.G. Cleary, A.J. McGregor, M.W. Pearson [5], but their work was based on identifying VoIP restricted to one protocol i.e H.323. It identifies VoIP by recognizing the TCP setup phase of H.323 protocol and then analyzing the UDP data for identification of RTP stream. According to their findings, H.323 application starts by requesting two TCP ports 1503 and 1720 to be used for call setup and call control. An H.323 application that wishes to connect to another H.323 user will connect to that machine on both ports 1503 and 1720. After establishing the connection, the UDP ports are negotiated for data transfer. Their identification process comprises of first detecting pairs of IP addresses communicating on TCP ports 1503 and 1720, and then identify that the UDP data between these ends.

Another study done by Tsutomu, Takayuki, Toshiya, Hideaki [6] is more QOS specific as it states “To provide a dependable VoIP service, it is necessary to apply traffic controls such as rate control and filtering by accurately
identifying the legitimate VoIP traffic from the prohibited traffic. Their technique is based on analyzing the packet exchange patterns including the signaling messages exchange (flow) and media data exchange (interaction) contrary to other methods of looking for specific port numbers and signatures. However their technique is restricted to Skype, Netmeeting, SIPphone VoIP applications and its practical implementation is yet to be realized.

The monitoring of standard and proprietary VoIP protocols was also carried out by Luca Deri [8]. He developed two open source applications named ntop and nprobe for this purpose. However Deri’s methods are again protocol specific.

“The Skype agent does not run on any standard source port. Skype randomly selects a source port for the agent to run on, and then communicates via either TCP or UDP, or both. The choice of the protocol that Skype uses depends on whether the agent is behind a proxy/NAT or has a public IP address. The destination IP addresses are not the same every time Skype runs and the destination port numbers are also not standard. All communication via Skype is encrypted. This also means that phone numbers called (Skype Out) or other data are also encrypted. In many cases, there is no direct communication between end users in Skype. All communication passes through intermediate nodes, and these nodes may be different for every call. Skype is a peer-to-peer protocol, which means that the peers (IP addresses) to which a Skype agent connects are many and the network is very dynamic, so these peers (and thus their IP addresses) keep changing. Skype provides voice, chat, file transfer and video services. It appears that all of these services are passed together, making it difficult to separate out voice, from chat, from video, etc.”

Detection of skype was facilitated by Baset [10]. His detection method is based on analyzing the packet exchange patterns; however his work is limited to older version of Skype. Elhert [11], Chun Ming [12] and Biondi [13] have all studied the Skype protocol. All their work is limited to Skype detection, and these methods cannot be used to for detection of other VoIP protocols.

4. Characteristics Analysis of VoIP Traffic

The captured traffic was analysed to establish the common characteristics of VoIP applications based on coherent packet attributes distinct to VoIP traffic. The evaluated factors of traffic include the time between first and last packet, the number of packets, average packet size, average packet/sec, total bytes, average bytes/sec and average Mbit/sec. The following traces, shown in figure 1, 2 and 3 respectively, reveal the above mentioned factors of various VoIP protocols including Skype, MSN, YAHOO and Google Talk.
When the packet characteristics of different applications were compared, it was noted that each application category has different discernable characteristics. The identification method then aimed at identifying these characteristics for voice conversations. For this purpose, all the available applications including file transfer, file downloading, multimedia sessions like video, game, music, instant messaging were compared to various VoIP services including MSN, Skype, YAHOO and Google talk. The summary of comparison of various characteristics of internet applications is given in table 1.

- The profound investigation of collected data revealed two distinguishing characteristics which can identify the VoIP from other network traffic irrespective of supporting protocols. It can be clearly seen from the table 1 that the packet size and packet rate are astonishingly noticeable in VoIP when compared to other applications. Based on these findings two observations were made which are elaborated in forthcoming paragraphs.
Observation No 1: Average packets / Sec rate is greater in VOIP as compared to other applications

An interesting characteristic of VoIP traffic noted was that voice traffic was found to have greater average packets per second rate as compared to other Internet traffic. The pie chart in figure 5 shows three VoIP services, each having a distinctively larger number of packets per second rate, whereas other traffic types have an almost same rate of packet transfer.

The graphs shown in figure 6, 7 and 8 shows the average packets/sec rate for Yahoo, Skype and MSN respectively. As shown in these graphs, the packet rate was found to be 21, 27.5, 25 and 35 respectively. The graphs shown in figure 9, 10 and 11 are the packets/sec rate of other Internet applications such as file sharing, games and video. Their packets/rate was found to be far less, lying between 2 – 9 packets/sec.
The graphs indicating the packet/sec for the well-known file sharing application Bittorent, game and video applications respectively are shown below.

Observation No 2: Average packet size in bytes is small in VOIP as compared to other applications

Another distinct feature found during the research work was that the average packet size of VoIP applications was much smaller than the packet size of other Internet applications. The pie chart shown in figure 12 shows the three VoIP services have minimal sized packets whereas other types of traffic use noticeably larger sized packets.

A study of the average packet size of the Skype protocol revealed that the packet payload size is smaller. An interesting feature about the Skype voice packet is that direction of the flow can be identified by packet size. For example, it was noted in one of the traffic capture where a Skype client, say SC-A, had placed a call to another Skype client, say SC-B, that all the packets being sent from SC-A to SC-B had a payload size ranging between 160-170. In reply to these messages, packets from SC-B to SC-A, the size of the packets fell between the 175-190.
Whatever the direction, however, the average packet size of Skype packets fell between 140-160 bytes. Average packet sizes in bytes for Yahoo, Google talk, MSN and Skype are 125, 178, 105 and 166 respectively whereas for rest of the applications it is well in hundreds between 400-800 bytes.

The small range of the packet payload size of Skype voice packets can be attributed to its encryption of the voice conversation. All Skype traffic is encrypted, making it difficult to discern any information about the voice conversation taking place using traffic analysis.

**Comparison of VoIP and file transfer applications**

As previously mentioned, differentiating traffic generated by VoIP applications is not a trivial task, however, using these finding identifying voice packets has become much simpler. When considering the case of file transferring application, the task is a bit simpler because most file sharing applications are based on the P2P network setup. In these applications, one node hosts a file, and the other node downloads the file, or part of the file, from that node. This results in most of the traffic being directed in one direction. The packet size of file sharing application is significantly larger than that of a VoIP voice packet. The average packet size of a file sharing application varies on application being used, but is well into the hundreds. VoIP packets on the other hand are purposely kept small to enable their faster transfer on the Internet. File sharing applications depend more upon transfer of the correct data as compared to the transferring time. Hence, due to the large size of the packets, the packet/sec rate is very slow.

**5. Generic VoIP Detection Algorithm**

The proposed detection algorithm for the NID implementation is given below.

**LEGEND**

<table>
<thead>
<tr>
<th>UI</th>
<th>Count of VOIP UDP Packets/second</th>
</tr>
</thead>
<tbody>
<tr>
<td>PT</td>
<td>Type of Packet</td>
</tr>
<tr>
<td>PL</td>
<td>Length of Packet</td>
</tr>
<tr>
<td>IT</td>
<td>Initial Time in seconds</td>
</tr>
<tr>
<td>FT</td>
<td>Final Time in seconds</td>
</tr>
</tbody>
</table>

**ALGORITHM**

1: Set UI = 0
2: Scan for incoming packets.
3: If (PT == UDP) // Filter UDP Packets
   3.1 If (PL >= 100 && PL<= 200) // Filter UDP Packets between 100 and 200 bytes
      3.1.1 If (UI == 0) // Set Initial Time Value
           IT = CurrentTime
      3.1.2 Else
           FT = CurrentTime
           If ((FT – IT = = 1) && (UI >= 20) &&
                (UI <= 40)) // Check if packet rate per second is between 20 and 40
                VOIP Activity Detected
           Set UI = 0
      3.1.3 UI ++

The algorithm captures all the packets coming on the wire and keeps looking for the UDP packets fulfilling following two conditions:-

(i) Average Packets/sec are between 20and 40.
(ii) Average packet size in bytes is between 100 and 200.

Whenever these two conditions are passed by the traffic NID application prompts with the VoIP activity detection.
6. Conclusion

Illegal VoIP call termination has become a souring issue for many countries. Implementing a mechanism that will enable the government to monitor IP traffic and detect VoIP usage has clear law-enforcement as well as financial implications.

Previously, traffic identification technology identified applications based on the port numbers in packet headers or signatures of payloads, but identification was not always accurate because these attributes could be altered or concealed by applications. Furthermore, these features are likely to change with the arrival of newer versions. A more effective detection mechanism is proposed in this paper, which is based on flow-level characteristics, such as packet inter-arrival time, packet rate and packet lengths. Basing detection on these features is far more effective as these features cannot be changed by applications.

The two observations mentioned in this paper are common to any type of VoIP services irrespective of protocols being used. These results will prove promising in VoIP detection as compared to conventional port and protocol specific detection methods which are not of generic nature.

Our proposed methodology, describes an accurate VoIP identification technology based on an analysis of application traffic characteristics. This technology makes it possible to realize more dependable VoIP detection technology by identifying VoIP based on multiple characteristics of packets.

The objectives set for this research were successfully accomplished and it opens the door for future research on VoIP detection based on unaltered characteristics of application.

Acknowledgements

We would like to acknowledge Dr Fauzan Mirza of National Institute of Information Technology, Pakistan for his valuable guidance throughout this research work. Without his kind support this work would not have been possible.

References

Fauzia Idrees received the B.E. in Electronics Engineering in 1995 and is completing her M.S. in Information Security from National University of Science & Technology, Pakistan. Presently, she is serving in Pakistan Air Force since 1997. Her research interests include network security, Voice over internet protocol and crytography.

Uzma Aslam is currently completing her MS in Information Security from the National University of Science and Technology, Pakistan. She received her BS degree in the field of Computer Science in the year 2004. Her area of research interest include Voice over IP security and network security.