Detection of Encrypted Voice over IP (VoIP) Traffic

By
Muhammad Mazhar Ullah Rathore
2009-NUST-MS-CCS-09

Supervisor
Dr. Fauzan Mirza
NUST-SEECS

A thesis submitted in partial fulfillment of the requirements for the degree of Masters in Computer and Communication Security (MS CCS)

In
School of Electrical Engineering and Computer Science,
National University of Sciences and Technology (NUST),
Islamabad, Pakistan.

(July 2012)

Approval

It is certified that the contents and form of thesis entitled "**Detection of Encrypted VoIP Traffic**" submitted by **Muhammad Mazhar Ullah Rathore** have been found satisfactory for the requirement of the degree.

Advisor: **Dr. Fauzan Mirza**

Signature: _____
Date: _____

Committee Member1: **Dr. Adeel Baig**

Signature: _____
Date: _____

Committee Member2: **Dr. Junaid Qadir**

Signature: _____
Date: _____

Committee Member3: **Dr. Adnan Khalid Kaini**

Signature: _____
Date: _____

Abstract

VoIP usage is rapidly growing technology due to its cost effectiveness, dramatic functionality over the traditional telephone networks and its compatibility to public switched telephone network (PSTN). In most of Middle East and Asian countries, like Pakistan, the commercial usage of VoIP is prohibited by national telecom authorities and they incurs loss of millions of rupees per year due to commercial usage of VoIP. Internet service providers may also want to prioritize VoIP from their paid customers. So Internet service providers (ISPs) and telecommunication authorities of these countries are interested in detecting VoIP calls to either prioritize or block VoIP traffic. Different techniques have been proposed for detecting VoIP such as port-based techniques, signature-based techniques, pattern-based techniques, and statistical analysis-based techniques. Port-based techniques, signature-based techniques, pattern-based detection is specific to some of VoIP applications and protocols. For generic purpose, only statistical techniques are used for better results but existing statistical analysis-based techniques have some limitations and they could not provide more efficient and accurate solution to such organizations. In this thesis, we propose statistical analysis-based solution using threshold values of flow statistical parameters to detect the VoIP media (voice) flows. The solution is generic, efficient, accurate and real time (to some extent) and can detect encrypted, non-encrypted, and tunneled VoIP. It is independent from any VoIP application, protocol, security mechanism, or any tunneling mechanism and practically implementable at telecommunication authority or ISP gateway to either block or prioritize VoIP traffic.

The proposed system is evaluated by accuracy, efficiency, and scalability. It has 97.54% direct rate (DR) and .00015% false positive rate (FPR). It detects VoIP calls from any VoIP application or protocol with 6 seconds. We compare our system with existing systems by accuracy and by features. Our system has better results and more features and fulfills the need of telecom operators and ISPs for detecting VoIP.

Dedication

In the name of Allah, the Most Gracious and Most Merciful

To my Parents, Brothers, and Sisters for their love and support

Certificate of Originality

I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, nor material which to a substantial extent has been accepted for the award of any degree or diploma at NUST SEECS or at any other educational institute, except where due acknowledgment has been made in the thesis. Any contribution made to the research by others, with whom I have worked at NUST SEECS or elsewhere, is explicitly acknowledged in the thesis.

I also declare that the intellectual content of this thesis is the product of my own work, except for the assistance from others in the project's design and conception or in style, presentation and linguistics which has been acknowledged.

Author Name: **Muhammad Mazhar Ullah Rathore**

Signature: _____

Acknowledgments

First and foremost, I am immensely thankful to Almighty Allah for letting me pursue and fulfill my dreams. Nothing could have been possible without His blessings.

More than anything, I would like to thank my Parents and family members who financially supported me and encouraged me to complete this challenging task. Special thanks to my mom, dad, and my grand mom for their prayers. I dedicate this thesis to my belated brother, Mustafa, who encouraged me and supported me as a friend throughout my educational carrier.

My heartfelt thanks to everyone at WISNET lab, committee members and all others who contributed in any way towards the successful completion of this thesis. Especially I am thankful to Mohsin Sardar and Mohsin Kazmi for supporting me in tunnel configurations, Saeed Ullah for their guidance at each phase of thesis, Muneem and Tayyab for helping me to create simple VoIP setup.

I am immensely grateful to my uncles, prof. Dr. Habib Ahmed Rathore and Jamil Rathore, for always being appreciative of my hard work, for his unending love and support and for encouraging me during the hardest of times.

Finally, this thesis would not have been possible without the expert guidance of my esteemed advisor, Dr. Fauzan Mirza, who has been a great source of inspiration for me during these years of research. Not only has he been readily available for me but he always responded to any queries that I might have, read through my draft copies, listened to my moaning and complaining and supported me every step of the way.

Muhammad Mazhar Ullah Rathore

Table of Contents

List of Abbreviations

VoIP	Voice over Internet protocol
PSTN	Public switched telephony network
PTA	Pakistan telecommunication authority
PTCL	Pakistan telecommunication limited
ISP	Internet service provider
IANA	Internet assigned numbers authority
FLB	Flow level behavior
VoBB	Video over broadband
MGCP	Media gateway control protocol
AH	Authentication header
ESP	Encapsulation security payload
SPI	Security parameter index
S-IP	Source IP
D-IP	Destination IP
S-Port	Source port
D-Port	Destination port
Max-diff-time	Maximum difference time
TP	True positive
TN	True negative
FP	False positive
FN	False negative
DR	Direct rate
FPR	False positive rate
HFBA	Host and flow behavior analysis
PSD-PA	Packet size distribution and port association
SIP	Session initiation protocol
SSL	Socket layer security
TLS	Transport layer security
IPSec	IP security
SRTP	Secure Realtime transport protocol
RTCP	Realtime transport control protocol

List of Tables

List of Figures

Chapter 1

Introduction

Voice over Internet Protocol (VoIP) usage is increasing day by day due to its low cost and dramatic functionalities. Support for public switched telephone network (PSTN) is also provided in VoIP so that the user can talk to the non-IP-based telephone from VoIP phone (PC-to-Phone call). Main steps that are involved in VoIP call setup are signaling and media channel setup (voice transmission). At both steps, different protocols are used. The signaling is used to setup the call between two communicating parties. The media channel setup is the actual voice transmission channel between two parties after a successful signaling. SIP and H.323 are main signaling and RTP is main media transmission protocol of VoIP. Detection techniques can be applied on any of these two steps. Some techniques detect VoIP traffic by examining signaling traffic and others by examining media traffic. There are also some other techniques that examine both signaling and media traffic.

Detection of VoIP is important for both telecommunication authorities and ISPs for blocking or prioritizing VoIP. Use of complex encryption and tunneling mechanisms for VoIP makes detection very difficult. VoIP signaling and media transmission both may be encrypted or any one may only be encrypted. The media session may be encrypted by SRTP, SSL/TLS, IPSec, or propriety protocols. The signaling may be encrypted by SIPS, SSL/TLS, SMIME, IPSec, or propriety protocols. There may also be such scenario where the signaling is encrypted by SSL/TLS and the media transmission is encrypted by SRTP. There are lot of different possibilities for encryption. Different techniques exist to detect VoIP traffic. These techniques are divided into 4 basic classes i.e. port-based techniques, signature-based techniques, pattern-based techniques, and statistical analysis-based techniques. Each type of technique has some limitations. Now a days, some of the techniques are rarely used due to complex, confidential and secure privacy protocols and

new VoIP applications and technologies are going to be introduced day by day to protect the privacy of VoIP conversation. So it is very difficult to come to an ideal solution with 100% accuracy and efficiency while detecting VoIP. For generic purpose, only the statistical techniques are used for better results. For specific protocols or applications, we can use port-based, signature-based, or pattern-based techniques for efficiency. Signature-based techniques are not applicable for encrypted VoIP traffic. Port-based and pattern-based techniques are not generic and can't be used for tunneled VoIP. Statistical techniques are generic techniques but existing statistical techniques could not detect all types of encrypted and tunneled VoIP efficiently and with higher accuracy. So they could not provide the practical solution to telecommunication authorizes and ISPs.

In this thesis, we are going to propose and develop a generic, efficient, robust, and practically implementable system to detect encrypted, non-encrypted, or tunneled VoIP media (voice) flows from network traffic. The system uses statistical measures of the flow and compares them with threshold values. The threshold values are taken from detailed statistical analysis of both VoIP and non-VoIP traffic. The voice traffic of mostly used VoIP applications such as Gtalk, Skype, Yahoo, MSN, Asterisk PBX with Blink, Eyebeam, X-lite, Zfone etc. are statistically analyzed. The results show that our system is best choice for telecommunication authorities and ISPs to either block or prioritize the VoIP calls.

1.1 Motivation

Over the years VoIP applications have gained much significance. Many VoIP applications are peer to peer such as Skype, Gtalk, and Yahoo messenger while others such as Asterisk PBX, Zfone, Eyebeam, and Blink are used for commercial purposes. In some countries, use of VoIP for commercial purposes is prohibited as it incurs loss of millions of rupees to telecommunication authorities, so telecommunication authorities are really interested in detecting and blocking the commercial usage of VoIP. Moreover ISPs or other service providers may want to prioritize VoIP traffic for paid customer. So detection of VoIP traffic is important by two aspects; one for blocking or restricting commercial usage of VoIP, other for prioritizing it. When the VoIP technology is introduced, all VoIP traffic travels in plain form (unencrypted) so the detection was not a difficult task but now most of the commercial VoIP application use complex encryption and tunneling protocols to secure VoIP traffic, the detection becomes problematic. Multiple

solutions [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15] exist for detecting encrypted VoIP. But none of the technique provides real time and practically implementable solution to detect encrypted and tunneled VoIP traffic from any VoIP application or protocol. So there is a need of an accurate, generic, efficient, and real time solution for VoIP detection that can be practically implemented to prioritize and block VoIP calls. The solution should not be dependent on any application, protocol, security mechanism, or any tunneling mechanism used by VoIP.

1.2 Aims and objectives

The primary objective of this thesis is to design and develop a system that can identify encrypted and tunneled VoIP conversations by giving one-way or two-way traffic. The proposed system should be generic, efficient, real time (to some extent) and practically implementable. With the passage of time VoIP service providers and clients are increasingly switching over the use of encrypted VoIP traffic. It is thus critical for the telecommunication industry to identify encrypted VoIP traffic. Moreover there are also some ISPs who want to prioritize VoIP traffic for paid customers. So this project is aimed to improve the current scenario by deeply analyzing the traffic semantically, behaviorally and statistically and propose an efficient solution to identify encrypted VoIP conversations by giving one-way or two-way traffic. The VoIP can be encrypted by secure VoIP protocols (such as SRTP, ZRTP), by secure tunnels (such as SSL/TLS or IPSec), or by other propriety protocols (e.g. Skype). The main objectives of the thesis are:

1. Thoroughly survey of the existing encrypted VoIP detection techniques and searching out the limitations of these techniques.

2. Analysis of voice traffic generated by mostly used VoIP applications such as Skype, Gtalk, Yahoo, MSN, Zfone, X-lite, Eyebeam, Blink, Asterisks PBX etc.

3. Analysis and identification of encrypted VoIP traffic that uses secure VoIP protocols i.e. SRTP, ZRTP.

4. Analysis and identification of encrypted VoIP traffic, that uses commonly used secured tunnels i.e. IPSec and SSL/TLS.

5. Based on analysis propose a generic, efficient, accurate and real time (to some extent) and practically implementable statistical analysis-based

solution that can detect encrypted, non-encrypted, and tunneled VoIP. The solution should overcome the limitations of the existing techniques.

6. Testing the systems on offline captured dumps, traces, and sample data.
7. Evaluate the system.
8. Deploying and testing the system on the real environment.

1.3 Major contribution

In this thesis, we propose an efficient, robust, and practically implementable solution to detect all types of VoIP media (voice) flows with higher accuracy. It is generic solution for VoIP media flow detection that does not depend upon any VoIP application, protocol, and any encryption or tunneling mechanism. We critically analyze the existing solutions for VoIP detection and also find out the limitations in the existing systems and reasons, why these solutions are not feasible for telecommunication authorities and ISPs for VoIP calls detection. Mostly used VoIP applications and protocols are semantically, behaviorally, and statistically analyzed. Moreover non-VoIP applications' traffic which have higher packet rates such as YouTube, torrent, antivirus updates, FTP downloads, online TVs, online games etc. are also statistically analyzed. SSL and IPSec tunnels are analyzed semantically and the VoIP into SSL and IPSec is analyzed statistically. On the basis of all these analysis, we propose a solution for our problem statement. The proposed system is finally tested and evaluated on both offline datasets and real time traffic by accuracy, efficiency, scalability, and by comparing with existing techniques.

1.3.1 Problem statement

"Design a generic, robust, efficient, and practically implementable statistical analysis-based solution to detect encrypted, non-encrypted, and any kind of tunneled VoIP media flows using threshold values of flow statistical parameters by giving one-way traffic or two-way traffic"

1.3.2 Problem breakdown

The research to reach at an efficient solution will be conducted according to the following action items:

1. Literature survey and problem identification: This is the first step that covers the exact problem identification. Different articles and papers

from journals and conferences have been studied to gain the complete knowledge of existing solutions. The limitations in the previous solutions are identified and reasons why these solutions are not feasible for telecommunication authorities and ISPs for VoIP calls detection are also found.

2. Datasets collection: In this phase, different traces and datasets are collected. These datasets contain different VoIP and non-VoIP data. The traces are collected for analysis. The traces are collected from NUST SEECS WISNET lab, from PTA, PTCL gateway, from home users, and sample traces downloaded from Wireshark site [16] and tstat site [17].

3. Semantic, behavioral and statistical analysis: The captured traffic is semantically, behaviorally and statistical analyzed. The semantic analysis are performed to detect the RTP, SRTP, ZRTP, SSL, IPSec traffic for analysis, the behavioral analysis will be conducted to identify the particular common behavior of VoIP applications, and the statistical analysis are performed to identify and confirm the VoIP flows by thresholds. The C programming with Winpcap library is used for analysis.

4. Algorithm design and testing on datasets: Based on statistical analysis, design an efficient solution for our problem statement and then test on the offline captured traces and dumps.

5. Testing solution on real environment: When the proposed solution gives better results at offline dumps, it is tested at online environment. If there is any limitation and error, it is removed.

6. Results, comparisons and evaluation: Finally the results are compiled and comparison is made against the previously proposed techniques. Evaluation of the proposed system with respect to accuracy, efficiency and scalability is done

1.4 Thesis organization

The remainder of this thesis is structured as follows:

Chapter 2, provides basics of VoIP that is useful to understand rest of the thesis. Complete survey of existing techniques is also presented in this chapter. All the existing VoIP detection techniques, the work that has been done using these techniques, limitations, and comparisons among these techniques are made in this chapter. Moreover, it also provides the comparison

of existing statistical analysis-based VoIP detection techniques.

Detailed experiment analysis and the proposed system design and implementation are given in chapter 3. This chapter presents the findings of semantic and statistical analysis. It also includes the details of datasets and traces collection, analysis tools, and the analyzed VoIP applications and their versions. Signatures of SSL, IPSec, RTP, SRTP, and ZRTP are also presented in this chapter. Finally the findings of statistical analysis are given based on which the system is proposed. The propose system design includes pseudocode,flow charts, and discussion. It also includes the algorithm for detecting VoIP hidden in IP layer tunnels.

In chapter 4, the results, evaluations and comparisons are made. The evaluation is done with respect to accuracy, efficiency, and scalability.

Last chapter, chapter 5, concludes the thesis work and gives the future directions to carry out this research work.

Chapter 2

Background and related work

2.1 Introduction

The main steps that are involved in VoIP call setup are signaling and media channel setup (voice transmission). The detection techniques can be applied on any of these two steps. Some techniques detect VoIP traffic by examining signaling traffic and others detect VoIP by examining media traffic. There may also be a technique that examines both signaling and media traffic. In this chapter we provide basics of VoIP and moreover we review existing VoIP detection techniques and approaches and the recent work that has been done using these techniques. There are basically 5 types of techniques that are used to detect VoIP traffic flows i.e. port-based techniques, signature-based techniques, pattern-based techniques, statistical analysis-based techniques, and hybrid techniques. This chapter mainly focuses on the basic methodology, usage, advantages and disadvantages of each of these techniques, and the recent work that has been done using these approaches. Our main focus is on the work done using statistical analysis-based techniques. Moreover, in this chapter, we discusses both the encrypted and non-encrypted VoIP detection techniques. The rest of this chapter is organized as follows. Section 2.2 provides basic definition of VoIP. Section 2.3 discusses basic functions and protocols used by VoIP. Section 2.4 describes simple VoIP setup. Section 2.5 discusses different possibilities of VoIP encryption. Section 2.6 discusses the existing techniques and the work that has been done using these techniques in the domain. Comparison of these techniques is done in Section 2.7. Section 2.8 discusses the limitations of the existing techniques and the need of the industry for VoIP detection solution.

2.2 What is VoIP?

Voice over Internet protocol (VoIP) is a mechanism that sends voice/video over the IP based network. It is also sometimes called IP telephony, Internet Telephony or Voice over broadband (VoBB).More precisely; we can say that VoIP is a combination of Internet technologies, communication protocols and transmission technologies that together used to transmit voice or multimedia session over the IP networks.

2.3 Basic functions and protocols of VoIP

Main steps that are involved in VoIP call setup are signaling and media channel setup (voice transmission). Both of these steps involve the use of different protocols. The signaling is used to setup the call between two communicating parties. The media channel setup is the actual voice transmission channel between two parties after a successful signaling; it includes digitization of analogue voice signal, encoding, packetization and transmission of the VoIP packets over the packet switched network. At receiving side, the opposite steps are performed to receive voice. Different codecs are used to encode voice.

Basically there are three methods /tools used to communicate via VoIP.

1. VoIP telephone
2. Normal phone with a VoIP adopter
3. Using a computer with a VoIP software, speaker and microphone

There are three types of functions and protocols used to setup complete VoIP call: signaling, media transmission, and control transmission.

2.3.1 Signaling

Signaling is a mechanism that provides and establishes connection between two communicating parties. Client sends the requests to different servers to connect and/or get the information about the other communicating party for communicating voice. The mostly used signaling protocols are:

1. SIP
2. H.323

3. SAPV2
4. SDP

2.3.2 Media transmission

When signaling is successfully completed and connection is established between two VoIP communicating parties then the voice is now to be transmitted. RTP is the most widely used media transmission protocol for voice transmission.

2.3.3 Control transmission

This function controls the whole media session between the VoIP communicating parties. Control protocols are used to control the media session. Different types of control messages are transmitted. The mostly used control protocols for media session are:

1. RTCP
2. MGCP

2.4 Simple VoIP setup

Four main entities are involved in a simple VoIP setup i.e. caller, callee, proxy server, and registrar. Registrar contains the information of all the clients whether they are online or offline, and what are the IPs of the clients (caller and callee). Caller and callee must be registered with registrar before communication. The caller wants to communicate callee, so it sends INVITE request to proxy server. The proxy server does not have the information of callee such as IP address, so the proxy server sends lookup message to registrar to consult the information of callee. The registrar searches the callee information and sends the information back to proxy server. The proxy server sends INVITE request to callee's address. The callee receives the invitation and reply OK to caller via proxy server. When caller receives OK message, he sends ACK to callee via proxy server. When callee receives ACK, the P2P connection is established between caller and callee. This connection is called media channel. Now using this connection they can talk to each other without the interference of any intermediate server. A simple VoIP setup is shown in figure 2.1.

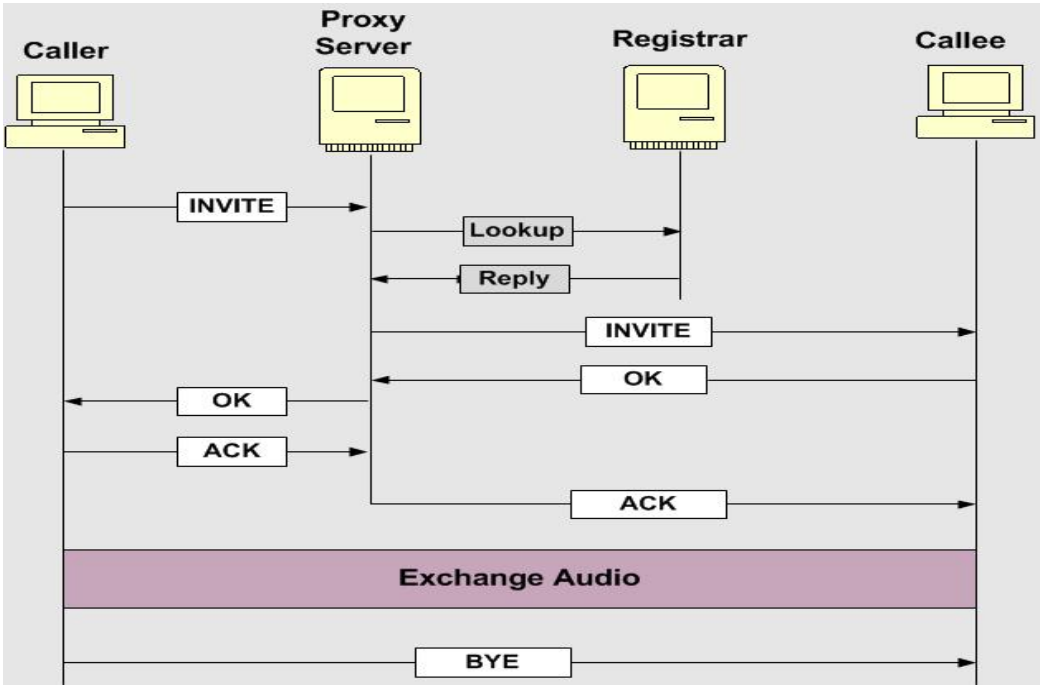

Figure 2.1: Simple VoIP setup

2.5 VoIP encryption mechanisms

There are different possibilities of encrypted VoIP. The signaling, media transmission both may be encrypted or any one of them is encrypted. The media session may be encrypted by ZRTP, SRTP, SSL/TLS, IPSec, propriety protocols etc. The signaling may be encrypted by SIPS, SSL/TLS, SMIME, IPSec, propriety protocols etc. There may be such scenario where the signaling is encrypted by SSL/TLS and the media transmission is encrypted by SRTP. There is lot of different possibilities for encryption. Different encryption algorithms are used for encryption. Our technique is not dependent on any encryption mechanism.

2.6 Overview of existing detection techniques

There are basically 5 types of techniques that are used for detecting VoIP traffic. These techniques are port-based techniques, signature-based techniques, pattern-based techniques, statistical analysis-based techniques and hybrid techniques. The details of each technique are presented in the next subsections.

2.6.1 Port-based techniques

Port-based analysis is easy to implement and fast but less accurate. By this technique, the traffic is classified only by examining the port number at the transport layer. IANA specified some standard ports to specific protocols such as VoIP uses 5060, 5061 ports for SIP signaling, 1718 to 1720 for H.323 signaling, and port 2427, 2944 for media gateway control protocol (MGCP), H.248 and Megaco protocols. SIP client uses ports 5060 to 5070 to communicate servers. So by port-based analysis if these ports are used at transport layer, the flow is detected as VoIP. Table 2.1 shows some of the standard ports for specific VoIP applications and protocols. In [18, 19], port-based analysis are used as helping information to detect VoIP. By [20], Skype VoIP traffic is detected by matching distinct Skype keywords, ports, and content. This technique is easy to implement and fast but it is not more accurate. Now non standard ports are used for every type of applications. Moreover the ports are dynamically allocated and in case of IP layer tunnels, the transport layer information is hidden. So in these cases this technique is useless and produces incorrect results.

Table 2.1: VoIP protocols standard ports

Protocol	**Default port**	**Transport protocol**
SIP	5060-5070	TCP/UDP
H.323	1718-1720	TCP/UDP
MGCP/Megaco/H.248	2427, 2944	TCP/UDP
Skype (client login)[20]	80, 443	TCP
Skype (authentication)[20]	33033	TCP

2.6.2 Signature-based techniques

Signature-based techniques overcome some of the drawbacks of port-based analysis. They detect VoIP using deep packet inspection by matching specific strings within packet payload. The VoIP protocols have distinct signatures that can be used for detecting VoIP for un-encrypted traffic. SIP packet has string “sip” within packet payload. RTP protocol has a distinct header that can be used to detect VoIP. RTP header mostly starts with 0x80, 0x81. Moreover ZRTP can also be detected by distinct signatures; it contains “1000xxxx5a525450” at the start of payload (ZRTP header). So these signatures are used for detecting VoIP. Moreover there are also some signatures that are specific to some VoIP applications. Some researchers analyzed these applications and found out the signatures specific to these VoIP applications.

Table 2.2: VoIP protocols and applications signatures

VoIP protocol	**Signatures**	**Place to find**
SIP	"sip"	application data
RTP/SRTP	0x80,0x81	RTP header, after transport layer header
ZRTP	"1000xxxx5a525450"	ZRTP header, Payload
Skype login [23]	"16 03 01 00 ** 42 cd ef e7 40 d7"	Payload, within transport layer packet
Skype (contents)[20]	/"getlatestversion?ver="	Payload, within transport layer packet

Signature-based techniques in [20, 21, 22] detect VoIP flows by VoIP application signatures. In [20], Skype VoIP traffic is detected by matching distinct Skype keywords, ports, and content. By [20] Skype packets sometimes contain the keyword "/getlatestversion?ver=" or "/getnewestversion" combined with a "/ui/" string. Moreover [20] points out that Skype extensively use port 33033 for TCP traffic, the outgoing data packets contain content "16 03 01 00 00", the incoming packets have content "17 03 01 00 00" and if the packet with these contents is blocked, the Skype tries to send a new packet that contains "03 01 00 cd 41 03 00 09 80 40 04 08 c0 00" and "00 0c 01 17 03 01 00". In [23], pattern-based as well as signature-based techniques are used to detect Skype traffic. By behavioral characteristics, [23] examines the packets for signature "16 03 01 00 ** 42 cd ef e7 40 d7" for Skype login. Table 2.2 shows signatures of different VoIP applications and protocols. Birke and Mellia proposed signature-based technique [24] for detecting RTP/RTCP over UDP detection by examining the RTP and RTCP headers. In [25] Skype traffic detection technique is proposed that is based on both signature as well as statistical analysis. By [25], Skype UDP ping carried out periodically by all Skype clients that consists keep-alive messages and the function field of the message is always 0x02. Similarly other signatures of Skype messages are identified.

The signature-based techniques are easy to implement, fast and efficient for un-encrypted data but these are useless for encrypted and tunneled data. In case of encryption, the data is totally senseless so these techniques could not find the VoIP signatures. Moreover signatures are changed from appli-

cation to application in case of proprietary protocols.

2.6.3 Pattern-based techniques

Pattern-based analysis depends upon signaling mechanism. These techniques are proposed to handle the shortcomings of port-based and signature-based analysis. By pattern analysis, the particular pattern of signaling communication of different VoIP applications is to be identified through which they provide connection between two VoIP clients to communicate voice. The way the VoIP client sends requests to VoIP servers to communicate with other client, is to be analyzed and then this particular way of signaling is used for detection. These techniques are powerful for the VoIP applications that use proprietary protocols for signaling i.e. Skype. Many researches [18, 19, 23, 26] have been done for detecting Skype traffic by pattern analysis.

In [18, 19] pattern-based analysis with port-based analysis are used to detect Skype traffic. In [18] forensic approach is used to investigate the Skype by deeply investigating the Skype communication and reveals the closed source Skype communication framework and detects Skype traffic. By investigating Skype, the paper [18] discusses 15 basic stages of Skype communication from start to end such as start up, registration; authentication, SN handshake etc. and also reveals all the entities and nodes that are participating in conversation such as Skype client, super node (SN), registration Skype node (RSN) etc. Moreover, for each stage the complete communication process and packets are analyzed for packet size, protocol and communication mechanism. Feng, Xiao and Zhi [23] use both port-based and pattern-based techniques. They analyze the Skype protocol with respect to its general and behavioral characteristics and use them to identify the Skype traffic and block it. By general characteristics, when Skype is being installed, the information of some of the super nodes, servers and login servers are stored in the local document "shared.xml". During the login process, Skype connects to one of the host in "shared.xml" list. After initial connection, the client retrieves the supper nodes information by sending one UDP and then TCP packets. One super node establishes a TCP connection to maintain connection between Skype client and Skype network. If network blocks TCP, the Skype can't login. TCP is used to authenticate and login user name and password.

Pattern-based techniques are good in some cases to detect encrypted VoIP but they are dependent on specific VoIP application. The signaling mechanism may vary from application to application, so in such cases they are not so efficient and accurate. Moreover these techniques are useless in case of IP

layer tunnels.

2.6.4 Statistical analysis-based techniques

To overcome the limitations of above discussed techniques, the statistical approaches came into research. By statistical techniques, some statistical measures are taken on flow features such as mean, standard deviation (S.D) of packet sizes and the packet arrival time measures are used for VoIP detection. Statistical analysis mostly performed on voice data but it can also be performed on signaling data. Statistical analysis can use different classifier to classify VoIP by taking flow statistics as input.

In [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 27] statistical approaches are proposed to detect VoIP. In [2], the IP addresses and ports are examined. In host behavioral analysis, the difference D between source ports and destination ports for a particular flow must be less than threshold. Moreover the inter-arrival packet time measure is used for detecting VoIP flows. Fauzia and Uzma [3] proposed a generic technique to detect the VoIP traffic generated by different VoIP protocols. They perform some statistical analysis on the traffic and separate out the VoIP media traffic by using traffic features that are difficult to alter such as packet interval time, packet sizes, rate of exchange. Freire and Ziviani describe a scheme [4] that detects the VoIP calls hidden in web traffic such as Gtalk and Skype traffic that uses port 80 and 443. Analysis are applied on media traffic by taking parameters such as web request size, web response size, inter arrival time between requests, no. of requests per page, page retrieval time. They use goodness-of-fitness test, the Kolmogorov-Smirnov (KS) distance and chi-square values and obtain metrics to identify the VoIP in web traffic. The scheme considers the key characteristics of normal behavior of web traffic (HTTP, HTTPS) and matched it to the actual traffic to identify VoIP. Yildirim and Radcliffe proposed statistical technique [8] to identify VoIP protocol within encrypted tunnel. They use probabilistic information of traffic to identify application protocols in tunnels. Their decision algorithm does Packet size distribution on packets that lies with a specified size rang. Ying-Dar and Chun-NanLu [9] also proposed a generic technique to classify the network traffic into different application types. They use packet size distribution (PSD) and assume that each application has a distinct PSD. They also use the port association techniques while classifying traffic by which if a port is consecutive to the previously identified flow port then it is detected as the part of the previous flow. Riyad [15] detects the VoIP traffic by using flow features, such as size and time and evaluates the three different machine learning (ML) techniques

Table 2.3: Modern statistical techniques

Ref.	**Year**	**Parameters used**	**Techniques used**	**VoIP applications tested**
[2]	2010	No.of ports, packet time	Difference of no. of ports, ratio of small and large inter-packet arrival times	Skype
[8]	2010	Packet size, packet time	Prob. density function, Packet size distribution (PSD)	Skype
[5]	2010	Packet size	packet size rang only	Own VoIP setup
[15]	2010	Packet size, time, mean, S.D, max-time etc.	C4.5, adaBoost, SBB-GB classifiers	Gtalk, Skype
[9]	2009	Packet size, ports	PSD, ports	Skype, MSN
[6]	2009	Packet size, time, direction	K-means classifier (only 1st few pkts)	Nil
[3]	2008	Packet size, exchange rate	mean and packet rate by threshold	Skype, MSN, Yahoo, Gtalk
[4]	2008	Request and respond size, time, no. of requests	Goodness-to-fitness test, KS distance, chi-square	Skype, Gtalk
[12]	2007	Packet size, time, flow duration	J48, REP tree	MSN, Skype
[27]	2006	Packet size, time	Flow level behavior (FLB)	SIPSoftphone, Netmeeting, Skype, Kaza

namely C4.5, AdaBoost, and SBB-GP for classifying Traffic. Toshiya Okabe proposed Flow level behavior (FLB) VoIP detection technique in [27] in 2006. This technique uses the packet size and inter arrival time to measure average of no. of packets per second, median and distribution of packet sizes for VoIP detection. Yildirim and Radcliffe [5] proposed a simplest statistical technique for VoIP identification that uses IPSec tunnel. It just considers the packet sizes of the traffic to detect IPSec tunneled VoIP. Table 2.3 shows the work done by using statistical techniques, the features and parameters, techniques used, and the VoIP applications on which the system is tested.

Statistical approaches are good and produce better results in case of encrypted VoIP. The results of statistical approaches are better than other approaches on latest VoIP applications but still the existing statistical techniques are not so efficient for IP layer tunneled VoIP detection. Moreover

most of the statistical approaches are not real time and need prior captured traffic to analyze. So these systems could not be practically implemented to block or prioritize VoIP efficiently and accurately with best results.

2.6.5 Hybrid techniques

The hybrid techniques use more than one of the above discussed techniques. More than one techniques are used to gain the advantages of multiple approaches and overcome the limitations of these techniques. The hybrid techniques produce better results than individual technique. In [18, 19, 24, 25], hybrid techniques are proposed for VoIP detection. Pattern-based analysis with port-based analysis are used in [18, 19] for Skype traffic detection. In [25], both signature-based and statistical approaches are used to detect Skype VoIP. The proposed technique in [25] also use pattern-based as well as port-based analysis. It analyzed Skype for Skype UDP ping, Skype UDP probe, Skype TCP handshake, and Skype authentication both statistically and behaviorally. In [24] both RTP/RTCP signature and statistical measures are used to detect VoIP.

2.7 Techniques comparisons

Port-based techniques are easy to implement and fast but these are not as much accurate as the need of the era. The usage of non standard ports makes port-based detection inaccurate. Moreover the ports are dynamically allocated and in case of IP layer tunnels, the transport layer information is hidden. So in these cases, port-based techniques are useless and produce inaccurate results. The signature-based techniques are easy to implement, fast and efficient for un-encrypted data but it is useless for encrypted and tunneled data. In case of encryption the data is totally senseless to understand so these techniques could not find the proper VoIP signatures. Moreover signatures are changed from application to application in case of proprietary protocols. Pattern-based techniques are good in some cases to detect encrypted VoIP but they are dependent on specific VoIP application. The signaling mechanism may vary from application to application, so in such cases they are not so efficient and accurate. Moreover such techniques are useless in case of IP layer tunnels. Statistical approach is a generic technique and produces better results in case of encrypted VoIP. The results of statistical approaches are better than other approaches on latest VoIP applications but they are costly in terms of speed because statistical approaches need some prior data. Table 2.4 shows comparison of these techniques in terms of scalability, efficiency,

encryption support, and part on which the technique is applied.

Table 2.4: Comparison of VoIP detection techniques

Technique	**Applied on?**	**Scalability**	**Performance (speed)**	**Encryption support**
Port-based	Signaling, voice data	Specific to application and protocol	Better	Yes, other than IP layer tunnels
Signature-based	Signaling, voice data	Specific to application and protocol	Better	No. (yes only for SRTP)
Pattern-based	Signaling	Specific to application	Good	Limited
Statistical analysis-based	Mostly on voice data	Generic	Bad	Yes

2.8 What's the need?

Each VoIP detection technique has some limitations. Now a days, some techniques are rarely used because of complex, confidential and secure protocols such as SSL and IPSec. In case of encrypted tunneled VoIP such as IPSec and SSL VoIP, the port-based and signature-based techniques are useless. The pattern-based techniques only provide limited detection for tunneled VoIP. Moreover these techniques are specific to VoIP applications or protocols. The statistical analysis-based techniques are better than all of other techniques. The statistical analysis-based techniques are generic techniques that can handle encrypted and tunneled VoIP. Only drawback of statistical analysis is that they need some prior data for statistical measures, so they are not so efficient in terms of speed.

The statistical analysis is the only way to cope with the needs of communication authorities and ISPs for detecting VoIP traffic but all the previously proposed statistical analysis-based solutions have some limitations. They could not be practically implementable for detecting all types of VoIP traffic efficiently. The statistical technique proposed in [2] could not provide a real time solution for VoIP detection; as first you have to calculate number of source and destination ports used for a particular flow. It has false positive ratio more than 10% which is still large. Moreover it could not handle VoIP

into IPSec. The technique proposed by Fouzia and Uzma [3] only considers UDP traffic. Some VoIP applications may use TCP when UDP is blocked. Moreover the SSL/TLS may also use TCP. There are many VoIP applications that transmit voice packets of less than 100 bytes, so in this case this statistical technique could not detect VoIP. It could not handle IPSec VoIP. Freire's technique [4] is not applicable for all types of VoIP. It is specific to VoIP hidden in the web traffic using port 80,443. It only supports the http version 1.1. Moreover it is also not real time detection and need more prior data for analysis. By this statistical technique, the anomalies and intrusions that use port 80,443 are detected as VoIP. It could not handled IPSec VoIP. The statistical technique proposed in [8] is not specific to VoIP rather it is a generic technique to identify any application protocol so the results in case of VoIP are not good. Moreover, no practically implementable algorithm is given to identify the VoIP in this paper. Only one VoIP application (Skype) is analyzed and tested. Only 3 voice codec schemes are analyzed by this technique. Moreover, only packet size is used to identify the VoIP traffic so more false positives. It could not handle VoIP that uses TCP. Ying-Dar and Chun-NanLu also proposed a generic technique [9] to classify the traffic but it is not specific to VoIP classification. It only analyzed two VoIP applications i.e. MSN and Skype. This technique only depends on packet sizes, so have more false positives and false negatives. Results show that in case of MSN VoIP detection, there is 9% false positive and in case of Skype VoIP detection, there is 18% false negative. Alshammari [15] analyzed only two VoIP applications (Skype, Gtalk) and built a VoIP detection solution. Other important application such as Yahoo, MSN, Zfone etc are given no importance. So the results are only specific to Gtalk and Skype. So it can't be used for blocking or prioritizing VoIP traffic from any VoIP application or protocol. Yildirim proposed a technique [5] for detecting IPSec VoIP but it only depends on packet size. So there is more false positive. No proper VoIP application is tested by this approach.

So there is a need of an accurate, generic, efficient, real time, and practically implementable statistical analysis-based solution that can detect encrypted, non-encrypted and tunneled VoIP. The detection algorithm should not be dependent on any VoIP application, protocol, security mechanism, or any tunneling mechanism.

Chapter 3

Experiment analysis and proposed system

3.1 Introduction

This chapter presents the complete analysis of Internet traffic especially the VoIP media traffic. What types of traces are collected, what are the environments and locations from where the datasets are collected, are presented in section 3.2. The tools that are used for analysis are described in section 3.3. Details of VoIP applications, that are analyzed, are presented in section 3.4. The tunneled traffic detection by signature is described in section 3.5. Moreover in section 3.6 the detailed statistical analysis are presented. On the basis of statistical analysis, the solution is proposed for detecting VoIP flows. Two main algorithms are proposed in this chapter. First algorithm detects encrypted, un-encrypted VoIP flows, or the VoIP flows that are hidden in transport layer tunnels. The other algorithm that detects VoIP flows hidden in IP layer tunnels such as IPSec is proposed in section 3.10. Moreover the solution is discussed in details and easily understandable by flow chart representation and by pseudo code. The solution is implemented and tested in C/C++ language.

3.2 Dataset collection

The datasets are collected at different time from different environments and locations for analysis and testing. The datasets contain different types of traffic that are collected from 1) NUST SEECS WISNET lab 2) home users 3) PTA and PTCL gateway 4) sample traces downloaded from Wireshark site [16] and 5) tstat Skype traces from tstat site [17]. The traffic traces are also

captured by making own simple encrypted and non-encrypted VoIP setups using Asterisk as VoIP server. These datasets are in TCPdump standard format. The list of collected traffic traces and datasets are given below:

1. Skype voice and signaling traffic. Some Skype voice samples are downloaded from Wireshark site [16] and tstat site [17].
2. MSN, Yahoo, Gtalk voice and signaling traffic.
3. ZRTP traffic from Zfone, X-lite applications.
4. Other VoIP messengers traffic such as Eyebeam, Blink.
5. The non-VoIP traffic traces of YouTube, torrents, antivirus updates, videos, online live TVs, audio songs, FTP downloads, and online games.
6. Mail servers traffic such as Gmail, Yahoo mail, Hotmail.
7. VoIP that uses SSL/TLS tunnel.
8. VoIP in other IP layer tunnel.
9. All mix traffic dataset that include both VoIP and non-VoIP traffic.
10. A simple VoIP setup is established using Asterisk as VoIP server which communicates voice between Eyebeam, Zfone, and X-lite clients using RTP and SRTP, ZRTP and SSL.
11. Moreover some sample traffic from different sites is collected that includes different protocols and applications data e.g. Bluetooth, chatting, DNS traffic, document retrieval, frame relay, remote access, SMTP, SSH, telnet-remote access traffic.
12. PTA and PTCL gateway datasets

All of these traces and datasets are statistically analyzed for VoIP detection solution and then tested on the proposed algorithm. Detailed information of these traces are presented in chapter 4

3.3 Analysis tools

Wireshark is used for capturing traffic. It is also used for simple analysis that can easily be done by this tool. C language is used for complex analysis. The code is written in C using Winpcap 4.1.2 library to analyze the offline as well as online traffic. The proposed algorithm is also developed in C using Winpcap in Visual studio 8.

3.4 VoIP applications

Different VoIP applications are analyzed and tested. The main VoIP applications whose media traffic is analyzed are Skype, Gtalk, Yahoo, MSN, Zfone, X-lite, and Asterisk server with Eyebeam and Blink as VoIP clients. Simple VoIP setup is established using Asterisk as VoIP server and Eyebeam, Blink, Zfone, and X-lite as clients. Different old and new versions of these VoIP applications are analyzed and tested. The traces of Gtalk beta version, Skype 4.0.0.215, Skype 5.5.0.119 and Skype 5.5.59.124, Yahoo 9.0.0.2152, Yahoo 10.0, Yahoo beta, Yahoo 11.0, MSN 7.0, MSN 8.5, 15.4.3538.0513 and Windows Live messenger are statistically analyzed and tested.

3.5 Detection of tunnels

The complete headers of RTP, ZRTP, SRTP, IPSec, and SSL/TLS protocols are analyzed to search out distinct signatures. RTP sends media traffic (voice) in plain form. In case of SRTP, the voice is transmitted in encrypted form but the header of SRTP is open. ZRTP provides the negotiation between two voice communicating parties (peers) for encrypted voice transmission and then voice is transmitted using SRTP. These protocols are working in between transport layer (UDP, TCP packets) and application layer. The headers of RTP and SRTP are almost similar except for some extra bits, so same signature can be used for detecting RTP and SRTP voice. It is also possible that RTP and SRTP is in GRE tunnel. The GRE tunnel can be detected from IP header information; the proto field of IP header contains 0x20 for GRE packets. The SRTP, RTP packets contain 0x80 at the start of RTP/SRTP header that shows the versions and padding bits etc. but it can also contain 0x81, 0x82 but they are used rarely. For analysis purposes, we only consider 0x80 as signature for RTP/SRTP and it gives satisfactory results.

IPSec and SSL tunnels are also analyzed. IPSec can be used in three forms, authentication header (AH) for authentication, encapsulation security payload(ESP) for encryption, or both of them. The proto field of IP header tells whether IPSec tunnel is used. The proto field contains 0x50 for ESP and 0x51 for AH. The packet payload contains security parameter index (SPI) as identifier of IPSec tunnel and ESP sequence number in IPSec packet in plain form and the other portion of packet is encrypted. So we can distinguish IPSec flows by source-IP, destination IP, and SPI (S-IP, D-IP, SPI). When ESP and AH both are used, the proto field of IP header contains 0x51

Table 3.1: Signatures of encrypted tunnels

Protocol	**Signature**	**Remarks**
RTP/SRTP	0x80,x81,0x82	Start of RTP/SRTP header
ZRTP	0x1000 1st two bytes 0x5a525450 5th-8th bytes	Start of ZRTP header
IPSec (AH)	0x51	Proto field of IP header
IPSec (ESP)	0x50	Proto field of IP header
IPSec (AH/ESP)	0x51 and 0x31	Proto field of IP header
SSL/TLS (alert)	ox15	Start of SSL header and also check version field
SSL/TLS (application data)	0x17	Start of SSL header and also check version field
SSL/TLS (change cipher)	ox14	Start of SSL header and also check version field
SSL/TLS (handshake)	0x16	Start of SSL header and also check version field
SSL/TLS (continuation data)	Check handshake and change cipher done?	If handshake and change cipher done and data is encrypted

(AH) and AH header which is in plain form contains value 0x30 for next header field.

SSL/TLS tunnel provides security at transport layer where the transport layer information (TCP/UDP) is open. Different types of messages are transmitted in SSL/TLS tunnel from negotiation to the end of communication i.e. handshake, alerts, application data, change cipher spec, and continuation data. The signatures of each type of message are different. Where the TCP header ends, the SSL header is started. At start of SSL/TLS header, the value of the content type is written of 1 byte. Content-type:0x15 represents that the payload contains alert, 0x17 represents that the payload contains application data, 0x16 represents that the payload is handshake that is just negotiation, 0x14 represents that the payload is change cipher spec that tells that the next communication will use the negotiated ciphers, so the communication will be encrypted after this packet. The next two bytes represents the version which is 3.0. The continuation data do not have any signature. If the handshake of SSL/TLS is done between two IP pairs and change cipher message is sent then the remaining communication will be continuation or application data if the packet is encrypted. We use the entropy measures to

find out whether the packet is encrypted or not.

The signature of SSL and other tunnels are given in table 3.1. All these tunnels are statistically analyzed for investigating statistical characteristics of VoIP.

3.6 Statistical analysis

Different types of VoIP and non-VoIP traffic is statistically analyzed to develop an efficient solution. Different VoIP and non-VoIP traces are captured. The statistical analysis are performed on traces by two ways; firstly, the statistical parameters are calculated and analyzed for each flow for complete session without considering time limit and in 2nd phase the statistical parameters are calculated and analyzed for each flow by taking 5,5 seconds chunks of traffic for each flow. We use IP layer (layer 3) packet size shortly refers as size and inter-arrival time as basic parameters for statistical analysis. All VoIP and non-VoIP applications are statistically analyzed in this way. Moreover the RTP, SRTP, ZRTP, SSL, TLS, IPSec protocols are deeply analyzed. The traffic of these protocols is detected by signatures and then analyzed from PTA and PTCL dumps.

Analysis of PTA, PTCL dumps show that VoIP applications that use SRTP and RTP mostly encode voice by G729, G723, and G711 codec. More than 60 % RTP conversation use G729 codec, 20 % use G723 and remaining 20 % use other codec schemes. These results are generated from PTA dump "dumpa" (2GB) collected in Jan 2011. The average packet size of RTP streams in this dump is 88 bytes (Wireshark analysis). The use of G729 codec is higher because G729 codec encodes 20 ms voice in a smaller size packet (i.e. 20 ms voice is encoded in 20 bytes). G711 encodes 20 ms voice in 160 bytes. G711 is a worst case of coding voice. Moreover in this dump out of 12353143 packets 6645638 packets are VoIP packets. So 53 % packets of PTA traffic are related to VoIP containing RTP, SRTP, RTCP data and 83% of UDP packets are VoIP packets. These dumps are also statistically analyzed. RTP traffic is detected from these dumps by signature and then the statistical parameter $\overline{X}$ (size) and S.D (size) in bytes for each flow is calculated and values ranges for RTP, SRTP flows are identified. The values ranges of statistical parameters of RTP/SRTP flows by considering the layer 3 header size in PTA, PTCL dumps are shown in table 3.2.

Different traces of Skype, Yahoo, Gtalk, MSN, Zfone, X-lite voices are

Table 3.2: Statistical measures of RTP/SRTP voice flows in PTA and PTCL dumps

Dump	$\overline{X}$ (size) (bytes)	**S.D** (size) (bytes)	**Remarks**
PTA-dumpa	59-200	1-70	
PTA-dump1	60-210	0-100	S.D 75-100 is rare
PTCL	60-218	0-60	

also analyzed by considering the traces as a whole and finding statistical parameters values for each flow. The values ranges of statistical parameters for VoIP media flows are found. The statistical parameters and the corresponding values ranges for all VoIP voice flows are shown in table 3.3. It is a rare chance that the value of S.D (size) is more than 75 bytes. We consider only IPs for distinguishing flows (due to tunneled voice) and RTCP and RTP packets from same source-destination are considered as one flow; so S.D (size) may be larger in few cases. In case of IP layer tunnels, only IPs are in plain form, all the other information is hidden so we distinguish flows only by considering IPs.

Table 3.3: Statistical parameters values ranges for VoIP application on packet sizes for voice flows

Trace	$\overline{X}$ (size)(bytes)	**S.D** (size)(bytes)	**Remarks**
Skype	60-130	0-27	
Gtalk	100-176	2-75	
MSN	120-140	7-61	
Yahoo	70-175	8-87	S.D above 75 is rare
Asterisk traces	190-214	0-65	
Zfone, X-lite	214-218	0-10	

The non-VoIP traces are also analyzed statistically. YouTube, antivirus updates, online live TVs, torrents, FTP downloads have higher data rates that can create confusion with VoIP flows but the flows of these application have higher packets sizes, $\overline{X}$ (size) and S.D (size) and also deviates from VoIP statistical parameters values ranges for other parameters. Shortly the values ranges for all parameters are different for both VoIP and Non-VoIP traffic. The packet sizes, $\overline{X}$ (size), and S.D (size) values for DNS are similar to VoIP but the packet rate is lower in this case. DNS has packet size of 181, 174 bytes for reply and 60, 70 bytes for request. Moreover IMAP and POP protocols are also analyzed, the packet lengths for these protocols is quite

higher i.e. above 200 bytes and mostly above 600 bytes.

The 2nd phase of analysis is most important. In this phase the flow parameters are examined by taking 5, 5, second traffic for each flow. Moreover the flow is distinguished by 4 tuples (S-IP, D-IP, S-Port, D-Port) and more statistical parameters are added for each flow to obtain efficiency and accuracy. All VoIP and non-VoIP traces are statistically analyzed by taking 5, 5 seconds traffic for each flow. The voice flows have distinct values for each statistical parameter. The main statistical parameters that are used to analyze each flow are:

- pkt-rate: Packet rate of the flow in packets/sec
- $\overline{X}$ (size) : Mean (average) of IP layer (layer 3) packets sizes of the flow in bytes
- S.D (size): Standard deviation of IP layer (layer 3) packets sizes of the flow in bytes
- Max-diff-time: Maximum difference between the current and previous packets' time for all packets of the flow in seconds
- $\overline{X}$ (diff-time): Mean (average) of the difference between the current and previous packets times in seconds
- S.D (diff-time): Standard deviation of the difference between the current and previous packets times of the flow in seconds

VoIP is not tolerant to delay, latency, jitter, and packet loss which affect the quality of voice. Voice packet total delay consists of packet creation time plus network transmission timeout plus receiving buffering and decoding time. Latency is the delay in packet delivery. Variation in delays is called jitter. More latency, jitter and packet loss degrade the quality of voice. The total delay of a voice packet is increasing function of packet size. If the packet size is larger, more voice is encoded in a packet; it will take more packet creation time, transmission time, and decoding time, resultantly the total packet delay is increased. To maintain the quality of voice, the delay should be bearable and the packet size should also be within a limit. Moreover in case of voice, the loss of large size packet means the loss of more voice which is not tolerant. Due to these facts the voice packet length must lie within a limit to maintain the quality. Jitter also affects the quality of voice as larger variation in packet delays does not produce clear voice at the receiver side. There should be bearable variation in delays in case of

voice packets to maintain voice quality. By considering these facts, we use IP layer packet size as the basic parameters for statistical analysis. Other P2P traffic is tolerant to delays and jitter. Hence we use $\overline{X}$ (size) and S.D (size) as statistical parameters for VoIP traffic analysis to distinguish voice flows. Moreover the ITU-T recommends to capsulate 20-30 ms voice in a packet for better performance and quality assurance. It shortens packet size and increases packet rate i.e. more packets per second as compare to other applications. So we also use packet rate as parameter for statistical analysis. It is also a fact that voice has a continuous behavior as the voice packets are continuously sent when a person speaks on VoIP phone. Very short time is elapsed between the current and previous voice packets. We consider this fact and take max-diff-time, $\overline{X}$ (diff-time), and S.D (diff-time) as statistical measures for time-based analysis to distinguish voice flows.

The main VoIP applications that are analyzed are Skype, Gtalk, MSN, Yahoo, Asterisk with Zfone, X-lite and Eyebeam. The parameters and the corresponding values ranges for voice flow for each VoIP application are shown in table 3.4. Pkt-rate value is mention in packets/sec, $\overline{X}$ and S.D values is in bytes and all other parameters values is in second. Other non-VoIP traffic such as YouTube, torrent, antivirus updates, FTP downloads, online live TVs, mail servers traffic (Gmail, Yahoo mail, Hotmail), online games traces are also analyzed on these parameters. The values of statistical parameters are quite distinctive for both VoIP and non-VoIP flows. So on the basis of these parameters we can identify VoIP flows efficiently and with accuracy.

Table 3.4: Statistical parameters values ranges for VoIP application considering 5,5 second traffic for each flow

Trace	**Pkt-rate**	$\overline{X}$ (size)	**S.D** (size)	**max-diff-time**	$\overline{X}$ (diff-time)	**S.D** (diff-time)	$\lvert\overline{X}$**-S.D**$\rvert$ (diff-time)
Skype	16-50	60-140	.38-27	.075-.393	.019-.061	.008-.12	0-.07
Gtalk	17-37	90-170	5-65	.101-.426	.027-.056	.011-.578	0-.02
Yahoo	12-37	64-170	1-75	.065-.49	.026-.086	.010-.073	0-.03
MSN	17-50	120-140	05-20	.06-.74	.020-.058	.005-.055	0-.02
Asterisk traces with Zfone, X-lite, Eyebeam clients	17-30	190-210	0-40	.02-.41	.010-.046	.05-.49	0-.032

Some VoIP applications like Gtalk, Skype, and Yahoo send small number of packets at start of the media session. The range of these packets is 2-15 packets in first 10-15 seconds. Moreover there is also sometime the case max-diff-time is quite high (i.e. greater than 1 sec) for first and/or last packet of the flow. In such cases the difference | $\overline{X}$ (diff-time)-S.D (diff-time) | also exceed from normal rang. Moreover sometimes during media transmission the ports are dynamically changed for media session but the new ports allocated lies within the pervious port percentile. On the basis of these detailed statistical analysis, we propose a solution for VoIP calls detection in section 3.7.

3.7 Proposed algorithm design

The proposed solution detects any kind of VoIP which is encrypted or non-encrypted by separating out the packets for each flow. The flow is distinguished by 4 tuples i.e. source IP, destination IP, source port, and destination port (S-IP, D-IP, S-port, D-port).

The flow with statistical parameters values is putted to the decision algorithm when the flow time reaches 5 seconds or the no. of packets reached 80 for each flow. In case of VoIP flow the number of packets is larger so it only considers that flow whose number of packets greater than 65 within 5 seconds to make decision for VoIP. The statistical parameters such as packet rate (pkt-rate), $\overline{X}$ (size), S.D (size), maximum difference time between current and previous packet (max-diff-time), $\overline{X}$ (diff-time) and S.D (diff-time) are calculated for the flow which has the number of packets greater than 65 within 5 seconds. 8 rules are investigated to make decision whether the flow is VoIP or non-VoIP. These rules are:

1. pkt-rate $>$ 13 packets/sec
2. $56 \leq \overline{X}$ (size) ≤ 210 bytes
3. $0 \leq$ S.D (size) ≤ 75 bytes
4. $\overline{X}$ (size) $\geq$ S.D (size)
5. $0 <$ max-diff-time $\leq .8$ seconds
6. $0 < \overline{X}$ (diff-time) $\leq .09$ seconds
7. $0 <$ S.D (diff-time) $\leq .25$ seconds

8. $0 < |\ \overline{X}\text{-S.D (diff-time)}\ | \leq .1$ seconds

These rules are obtained by detailed statistical analysis of different VoIP applications and codecs voice traffic for individual flows and also by studying different standards and facts of voice traffic . The decision is made on the basis of these 8 rules. A flow is a VoIP flow if and only if first 4 rules are true and at least 3 rules from last 4 rules are satisfied. A flow is confirmed non-VoIP if the rate rule is true but any one rule from rule 2, 3, 4 is false. If the flow is neither VoIP nor non-VoIP then it is either a suspected or not to be decided yet. If the flow is suspected for first 5 seconds traffic then it is reinvestigated for next 5 seconds traffic and if it remains suspected 3 times then it is detected as non-VoIP flow. The flow chart of the complete algorithm is shown in figure 3.2, figure 3.3, figure 3.4. The detail discussion is made in discussion section and the complete pseudo code is presented in section 3.9.

3.8 Discussion

The proposed algorithm detects and separates out the VoIP flows from all the network traffic. It classifies the flow as VoIP flow, non-VoIP flow and suspected flow. If a flow is suspected more than 3 times than it is declared as non-VoIP flow. The suspected flow is reinvestigated on next 5 second traffic. If suspected flow next time fulfils the VoIP flow characteristics then it is declared as VoIP and similarly if it meets the characteristics of non-VoIP then it is declared as non-VoIP flow. Moreover there may be such flows that are untreated because they are not assured to be VoIP or they are not yet to be decided and need to be processed next time.

The proposed algorithm separates out the flow by distinct source-IP, destination-IP, source-port, destination-port (S-IP, D-IP, S-port, D-port). The solution is divided into three main interlinked processes namely 1) flow registration and main decision process 2) VoIP flow detection process 3) non-VoIP flow detection process. First process does some parameters calculations and updations and makes decisions by calling the VoIP flow detection process. VoIP flow detection process detects whether the flow is VoIP or not and returns the result to flow registration and main decision process. It does some statistical calculations and checks rules and flow statistical measures. It also calls non-VoIP flow detection process to find out whether the process is non-VoIP or suspected. These processes are interlinked and viewed as in figure 3.1.

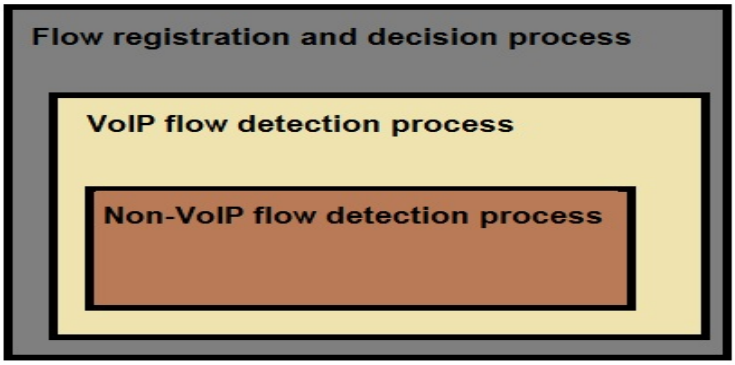

Figure 3.1: Process interrelation

The flow chart for the first process, flow registration and main decision process, is shown in figure 3.2. It separates out each flow and calculates and updates parameters values when a new packet comes. It sends the flow with corresponding statistical values of each parameter to VoIP flow detection process to detect whether the flow is VoIP or not. VoIP flow detection process investigates whether the flow is VoIP, non-VoIP or suspected flow and returns the result. It detects non-VoIP flows by calling non-VoIP detection process. Non-VoIP flow detection process investigates whether the coming flow is confirmed non-VoIP or need to be reinvestigate. The flow chart of process 2 and 3 are shown in figure 3.3 and figure 3.4 respectively. The functionality of each process is discussed in detail in next subsection.

3.8.1 Flow registration and main detection process

When a new packet is captured it is sent to flow registration and main decision process so that the particular flow should be identified for this packet. When the flow for which the packet belongs is identified, the calculations and decision process is started. This process checks whether the packet belongs to the previously registered flow or not. If it does not belong to one of the registered flows then it is registered as new flow uniquely identified by 4 tuples (S-IP, D-IP, S-port, D-port) and parameters values, calculated from packets size and time, are stored against the new flow. For the coming packet that belongs to one of the registered flow, the process checks whether the given flow has already been classified as VoIP or non-VoIP. The process does nothing for VoIP/non-VoIP classified flows when a new packet belonging to that flow comes. In such case the process will return to next packet. The flow that has not been detected as VoIP or non-VoIP, the process updates the value of flow statistical parameters by adding the new packets statistics

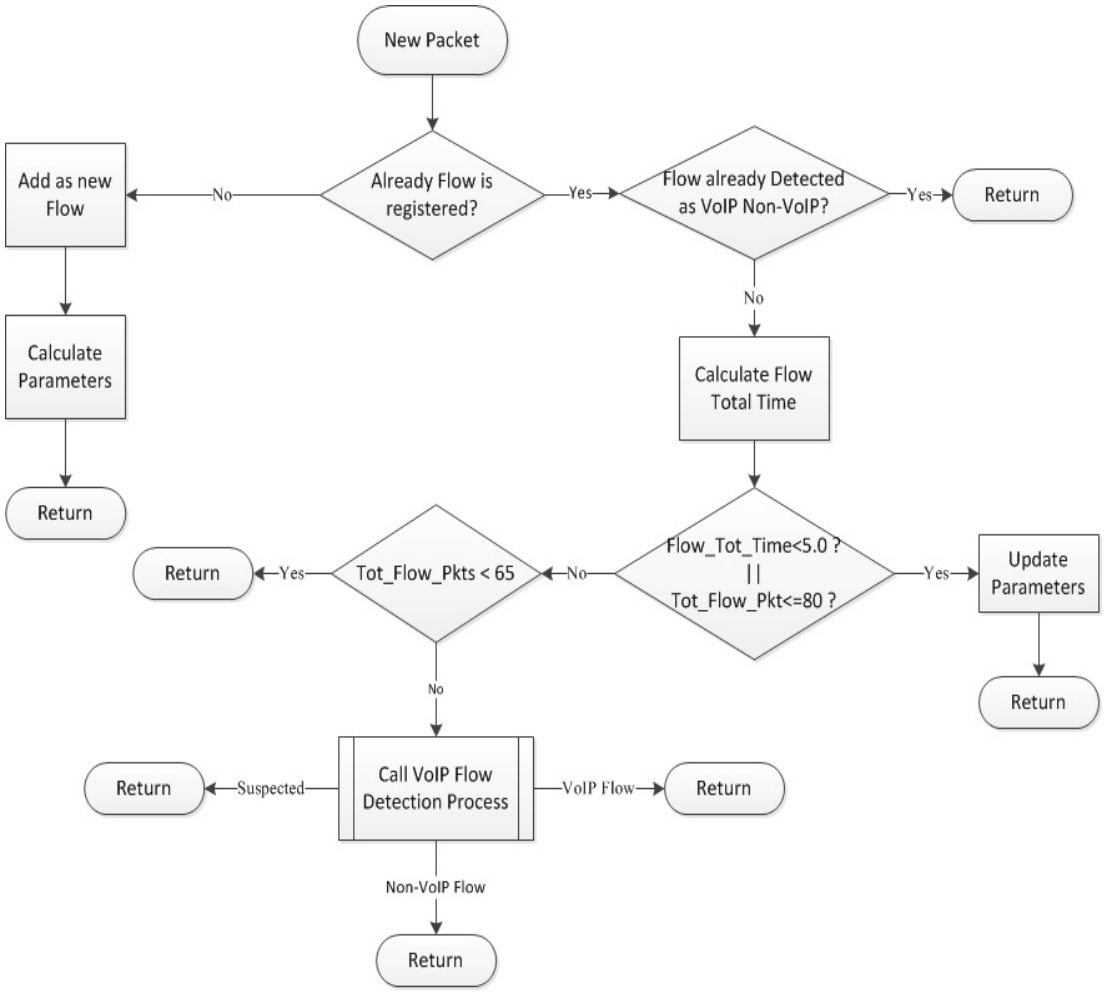

Figure 3.2: Flow registration and main decision process

until the flow elapse time reaches 5 second or the total number of packets for that flow reaches 80. When the time and number of packets condition satisfied for the flow, the process calls the VoIP detection process to classify the flow as VoIP or non-VoIP. The minimum number of packet for the flow to be VoIP is greater than 65 within 5 seconds time. The VoIP detection process returns whether the flow is VoIP, non-VoIP or suspected one.

3.8.2 VoIP flow detection process

The process takes a flow with statistical parameters as input and determines whether a flow is VoIP, non-VoIP or suspected flow. It uses non-VoIP detection process to determine whether the flow is confirmed non-VoIP so that it

can't be further investigated when a new packet belonging to that flow is captured. VoIP flow detection process calculates the statistical measurements from the flow parameters and checks them on predefined rules for VoIP, defined in section 3.7. The statistical measures are: $\overline{X}$ (size), S.D (size), pkt-rate, max-diff-time, $\overline{X}$ (diff-time), S.D (diff-time). For VoIP flow, first 4 rules must be true and at least 3 rules from last 4 rules must be satisfied. If less than 3 rules are satisfied from last 4 rules then the flow is declared as suspected flow and need to be further investigated for next 5 second traffic for this particular flow. The flow is non-VoIP if the flow is suspected more than 3 times. If any rule from first 4 rules is not satisfied then the non-VoIP flow detection process is called to confirm whether the flow is confirmed non-VoIP or suspected.The flow chart of VoIP flow detection process is shown in figure 3.3.

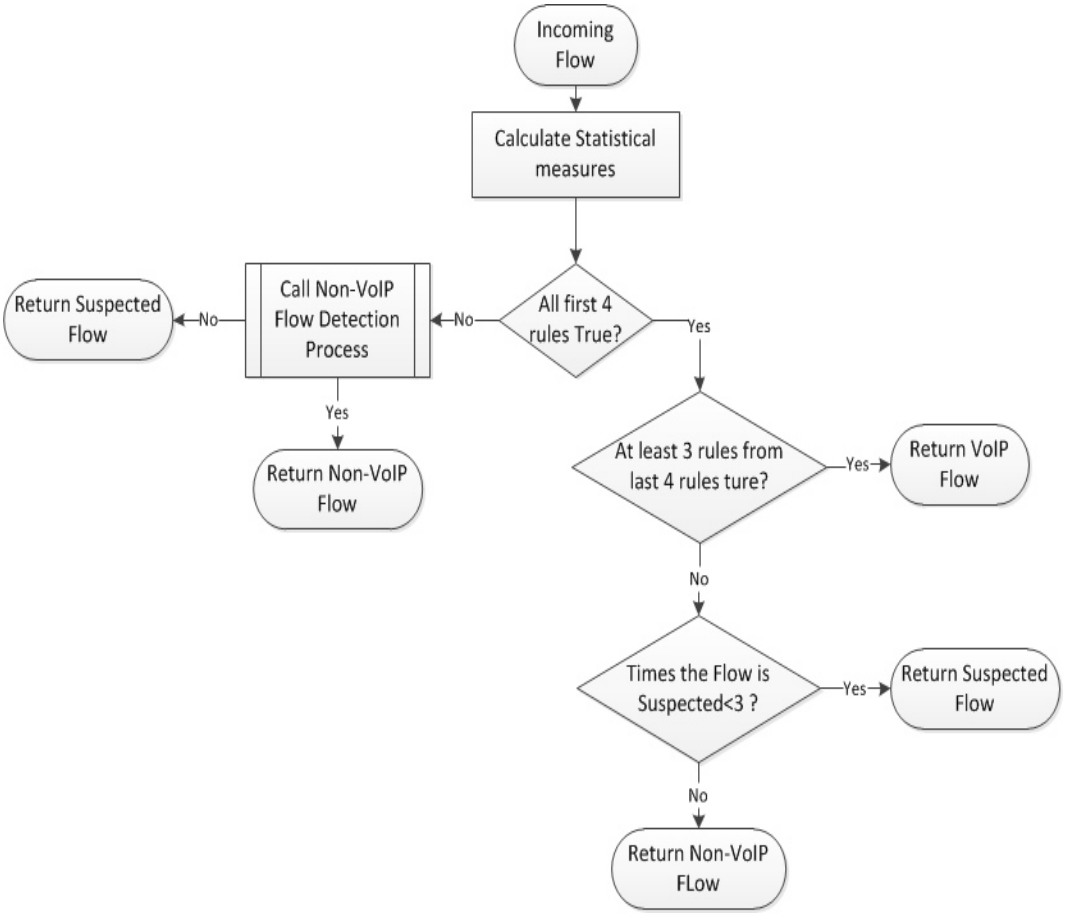

Figure 3.3: VoIP flow detection process

3.8.3 Non-VoIP flow detection process

This process takes the flow statistical measures as input and examined the flow for non-VoIP. The flow is confirmed non-VoIP depending on the rules described in section 3.7. It first checks whether the flow is suspected more than 3 times, if yes, it classifies the flow as confirmed non-VoIP flow. If the pkt-rate rule, rule 1, is true but any one rule from rule 2, 3, 4 is false then the flow is also detected as non-VoIP flow. Otherwise the flow is declared as suspected flow. The confirmed non-VoIP flow does not need to be reinvestigated when a new packet belonging to non-VoIP flow is captured. The flow chart of non-VoIP flow detection process is shown in figure 3.4.

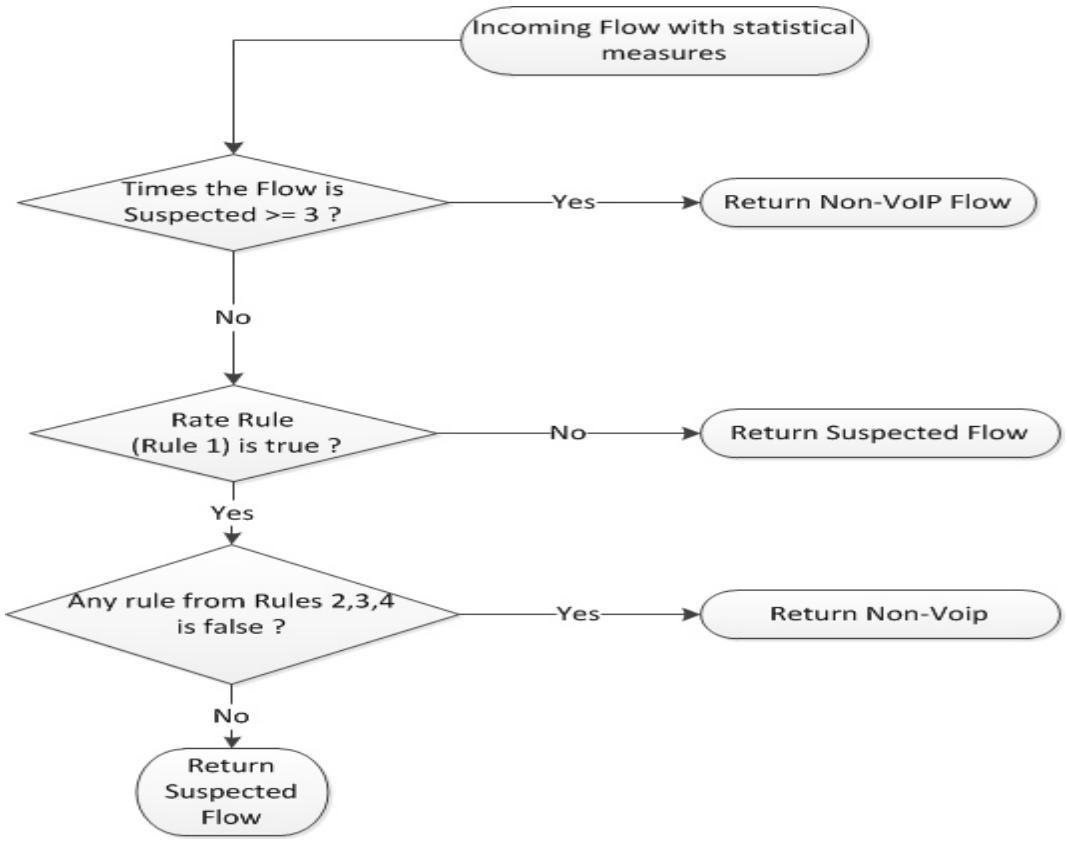

Figure 3.4: Non-VoIP flow detection process

3.9 Pseudo code

The algorithm use the rules defined in section 3.7. The pseudo code is:

1. Capture packets and determine the flow from which it belongs. If no flow found then register it as new flow uniquely identified by (S-IP, D-IP, S-port, D-port) and calculate parameters.

2. Capture the first 80 packets or all the packets within 5 seconds for each flow.

3. Investigate the flow for VoIP if the flow total packets greater than 65 within 5 seconds.

4. Check rules for each flow for VoIP detection

   (a) If (all first 4 rules == true) and if (at least 3 rule from rule 5, 6, 7, 8 == true) then the flow is VoIP flow.

   (b) If (all first 4 rules == true) and if (less than 3 rule from rule 5, 6, 7, 8 == true) then the flow is suspected. Suspected flows are reinvestigated for next phase (next 5 seconds traffic).

   (c) If (rate rule == true) and if (any rule from rule 2, 3, 4 == false) then the flow is confirmed non-VoIP.

   (d) If (flow suspected $\geq$ 3 time) then it is non-VoIP flow.

3.10 Proposed algorithm design for IP layer tunnels

The detection algorithm for IP layer tunneled (such as IPSec) VoIP is slightly different. For such type of tunnels, the transport layer information is hidden so we do not have the knowledge of ports. Only the IP layer information is in clear form. So we distinct each flow by IP addresses (S-IP, D-IP) only. In case of IPSec tunnel, we distinguish flows by (S-IP, D-IP, SPI). For each flow we calculate statistical parameters mentioned in section 3.6. The same detection algorithm is used for detection of VoIP flows hidden in IP layer tunnels with a slight modification. There are two main points that are different from the previous algorithm. These points are:

1. The IPs (S-IP, D-IP) only used to distinguish flows. (S-IP, D-IP, SPI) in case of IPSec.

2. Remove the size of IPSec headers from overall IP layer (layer 3) packet size while calculating statistical parameters related to size.

Chapter 4

Performance evaluation

4.1 Introduction

This chapter provides the complete results of the proposed system on different datasets, traces as well as on real time traffic. We evaluate our proposed system with respect to accuracy, efficiency, and scalability in terms of usage. At the end, the proposed system is compared with existing statistical VoIP detection techniques. The results are obtained by considering offline captured traffic, own VoIP setup traffic, tstat Skype traces, and real time traffic that contains VoIP and non-VoIP packets.

Table 4.1: Skype tstat traces

Trace	**codec**	**transport protocol**	**Size (MB)**	**Duration (sec)**
E2E-140606-1	G729	UDP	8	905
E2E-140606-2	iLBC	UDP	11	1003
E2E-140606-3	iSAC	UDP	12	1116
SkypeOut-260906-1	G729	TCP	9	919
SkypeOut-260906-2	G729	UDP	7	910
Internet-E2X		TCP	212	343562
Internet-E2O		UDP	264	343562
Internet-E2E		UDP	4GB	344700
Internet-SIG		UDP	6GB	344700

The system is implemented at real time environment for more than 5 hours. Voice conversation is done using different VoIP applications with different versions such as Skype 4.0.0.215,Skype 5.5.0.119, Skype 5.5.0.124, Yahoo 10, Yahoo 11, Yahoo 11.5, MSN Live messenger, Gtalk, and Gmail voice. Moreover widely used non-VoIP applications are also run for real

time system testing. Non-VoIP traces contain the traffic of emails, YouTube, LMS, online movies, songs, online games, online TVs, FTP uploading and downloading, torrents downloads, social networking sites access e.t.c.

Table 4.2: VoIP testing traces

VoIP application	**versions**	**Max-size (MB)**	**Max-duration (sec)**
Gtalk	1.0.0.104 beta, Gmail voice	3	504
Skype	4.0.0.215, 5.5.0.119, 5.5.0.124	3	664
MSN	7.5,8.0,15.4, Windows Live messenger	1	88
Yahoo (SSL tunnel)	9.0, 10.1, beta, 11.0	2	332
Mix VoIP		40	10714

The system is also testing on sample traces containing traffic of different protocols such as remote access, RDP,VNC, FTP, SIP, H.323, HTTP, HTTPS, Bluetooth, chatting, compression, DNS, TFTP, MSN file transfer, and web mails e.t.c. taken from wireshark site [16].

Table 4.3: VoIP setup testing traces

Trace	**Size (MB)**	**No. of files**	**Duration (sec)**
A-RTP-RTP	1.5	4	478
B-RTP-SRTP	.5	4	66
C-SRTP-RTP	.5	4	82
D-SRTP-SRTP	1.5	4	151
Zfone-X-lite	.25	1	32
Asterisk voice	4.5	1	151

Large set of traces is collected from different sources for testing purposes. Large size dumps are collected from PTA, PTCL gateway. These dumps are dumpa (2GB), dumpb (2GB), dump1 (4GB), and hundreds of PTCL dumps of 1GB. Some Skype traces are collected from tstat site [17] which are mentioned in table 4.1 with codec, transport protocols used, size of trace, and

time duration of the traffic. These traces contain both UDP and TCP voice conversations. Mostly used VoIP applications are also tested for detection. The information of the VoIP application traces which are tested are shown in table 4.2. Moreover a simple VoIP setup is also established for testing using Asterisk as VoIP server and Eyebeam and Blink as clients that communicate both encrypted and non-encrypted voice. Table 4.3 presents the information such as size, number of captured traces, and duration of the traces e.g. "C-SRTP-RTP" is a conversation of two VoIP clients in which one side traffic is encrypted by SRTP and other side traffic is un-encrypted and size of this traces is .5MB and duration is 82 seconds. non-VoIP traffic traces are also tested for checking how much flows are incorrectly identified as VoIP. The information of tested non-VoIP traces are shown in table 4.4.

Table 4.4: non-VoIP testing traces

Trace	**Max-size (MB)**	**No. of files**	**Max-duration (sec)**
Gmail-Yahoomail	3	5	156
Hotmail	3	3	101
Mix (VoIP-NonVOIP)	65	4	1023
NonVoIP-mix	112	6	1331
Torrent-YouTube-Gmail	1	1	431
YouTube	9	6	97
Online TV	2	1	88
Bittorent	150	5	2043

The results are obtained by considering all of these captured datasets as well as real time traffic. The results with respect to accuracy are discussed in section 4.2, the efficiency measures are discussed in section 4.3. The system is tested on dual core 2.3 GHz processor with 3GB RAM on a windows environment.

4.2 Accuracy

We use the typical parameters that are mostly used for measuring accuracy performance of the system. The parameters are true positive (TP), false negative (FN), true negative (TN), false positive (FP), direct rate (DR), and false positive rate (FPR). TP is the measure of flows that are correctly identified as VoIP flows. FN is the measure of flows that are incorrectly

identified as non-VoIP flows. TN is the measure of flows that are correctly identified as non-VoIP flows. FP is the measure of flows that are incorrectly identified as VoIP flows. DR and FPR measures also reflect the correctness of the system that are also used by [15]. DR reflects how much VoIP flows are correctly identified as VoIP flows. DR is calculated by equation 4.1. FPR measure reflects how much non-VoIP flows incorrectly identified as VoIP flows. FPR is calculated by equation 4.2. The ideal solution is that which has 100% DR and 0% FPR. So there should be highest possible DR and lowest possible FPR for more accurate system.

Table 4.5: Overall accuracy results

Traffic	**TP**	**FN**	**FP**	**TN**	**DR= TP/(TP+FN) %**	**FPR= FP/(FP+TN) %**
Real-time traffic	37	3	1	5000	92.5	.0002
Off-line traffic	56	2	2	15000	96.56	.00013
Own-VoIP-setup traces	16	0	-	-	100	–
Tstat Skype traces	882	20	-	-	97.78	-
Overall	991	25	3	20000	97.54	.00015

$$DR = TP/(TP + FN) \tag{4.1}$$

$$FPR = FP/(FP + TN) \tag{4.2}$$

Table 4.5 shows overall accuracy results considering real time traffic, offline captured traces, sample traffic, and tstat Skype traces. Our system has 97.54 % DR which is quite higher and .00015% FPR which is quite lower. Accuracy results on VoIP applications traces are shown in figure 4.1 Gtalk, MSN, Zfone, X-lite, and Asterisk with Eyebeam and Blink as clients has 100% TP and 0% FN. Only Skype and Yahoo has TP bit lower than 100%. Accuracy results on tstat Skype traces with respect to TP and FN are shown in figure 4.2. Only Skype trace "SKYPE-TCP-E2X" that have TCP voice communication has quite higher FN nearer to 17% but overall accuracy performance on these traces is better. Accuracy results are also obtain by considering the voice traffic of different codecs. G729, G711, iBLC, and iSAC codec traffic is tested. The results on these codecs are shown in figure 4.3. All these codecs have 100% TP.

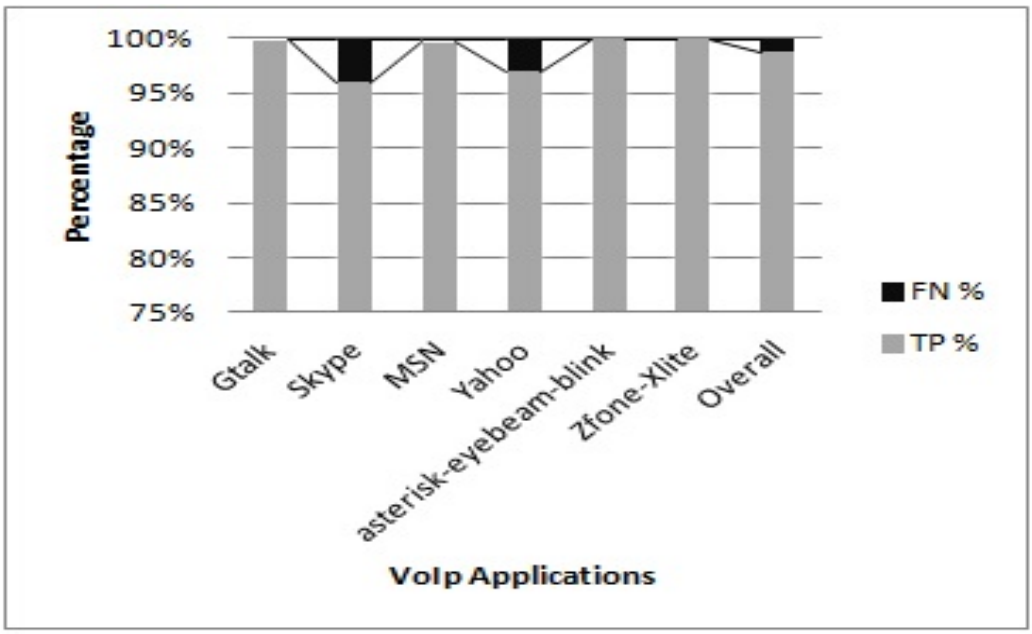

Figure 4.1: Accuracy results on VoIP applications

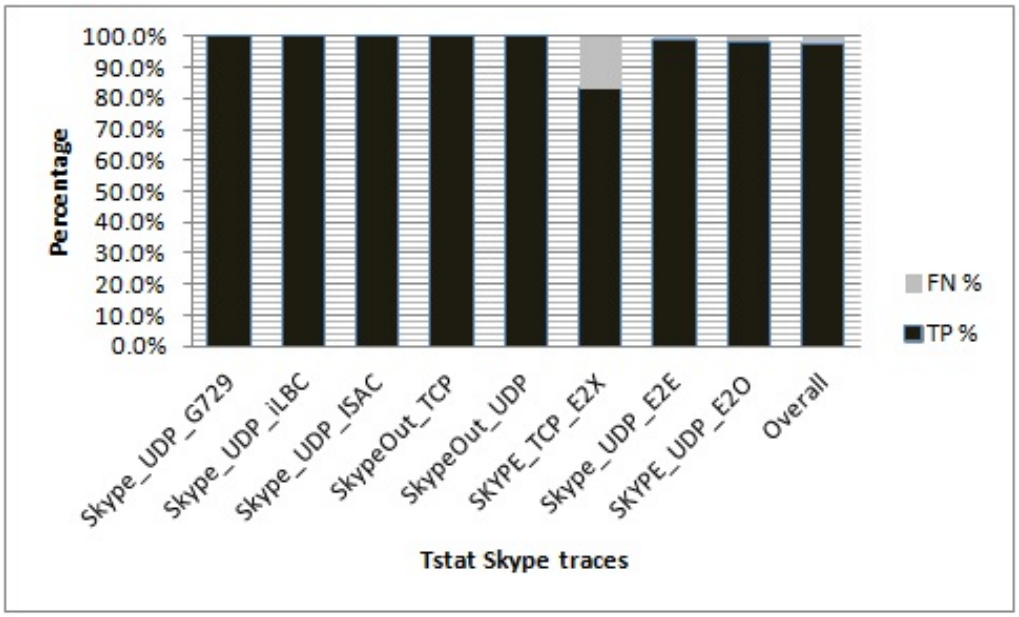

Figure 4.2: Accuracy results on tstat Skype traces

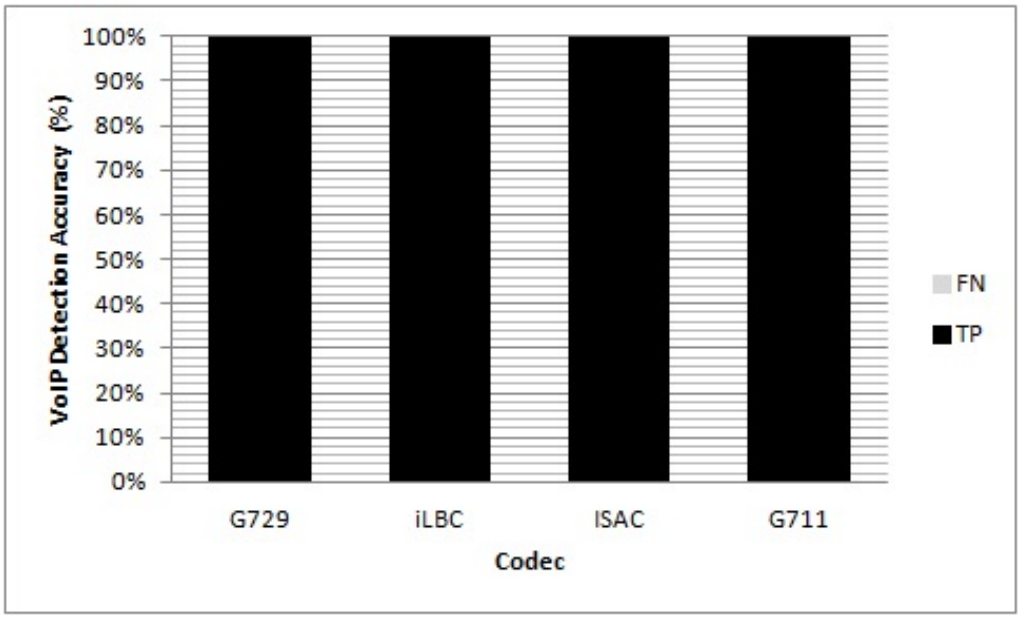

Figure 4.3: Accuracy results on different codecs

4.3 Efficiency

The efficiency is measured in terms of VoIP flow detection time, average number of packets processed by the system per second, and processing time on large size datasets. The average VoIP detection time for voice traffic is always less than 6 seconds. Figure 4.4 shows the average detection time of voice flows on different VoIP applications. These results are taken from real time implementation by communicating voices using different VoIP applications. MSN, Yahoo, and Skype voice flows are detected within 5 seconds and Gtalk, Gmail voice flows takes more than 5 seconds to be identified. Gtalk and Gmail voices takes more time to be identified because they send less numbers of packets as compare to other applications and system makes decisions when specific number of packets has been received or specific time has been passed for each flow.

Our implementation processed more than 10000 packets per seconds on defined large size datasets. Figure 4.5 shows the efficiency results in terms of average packet processed per second on different large datasets. Our implementation processed more than 20000 packets per second on 6GB tstat Skype trace.

The time consumed by our implementation on different large datasets is

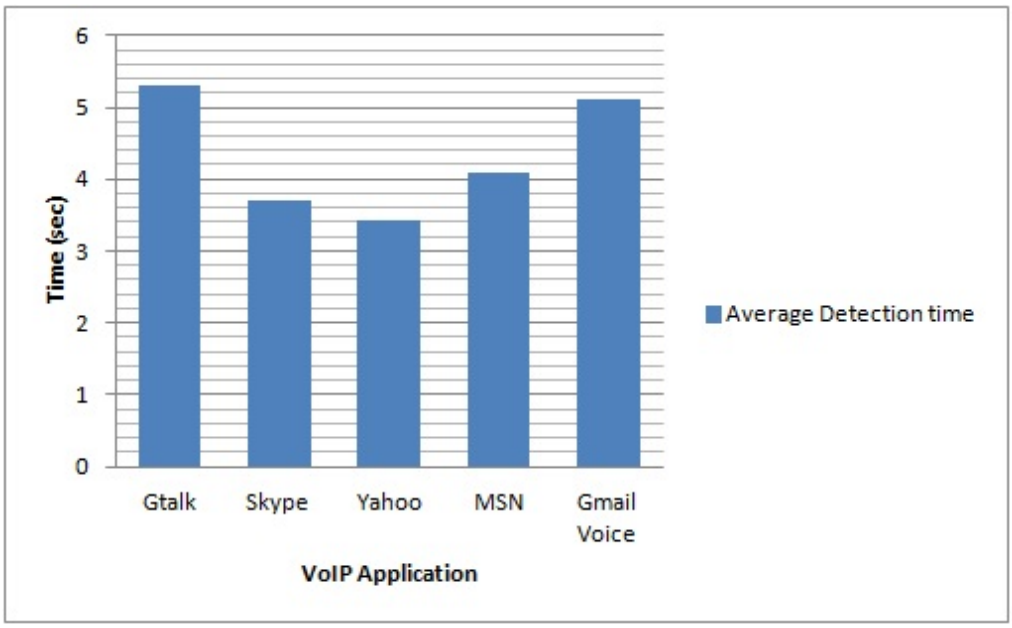

Figure 4.4: Voice flow detection time taken by the system on different VoIP applications voice flows

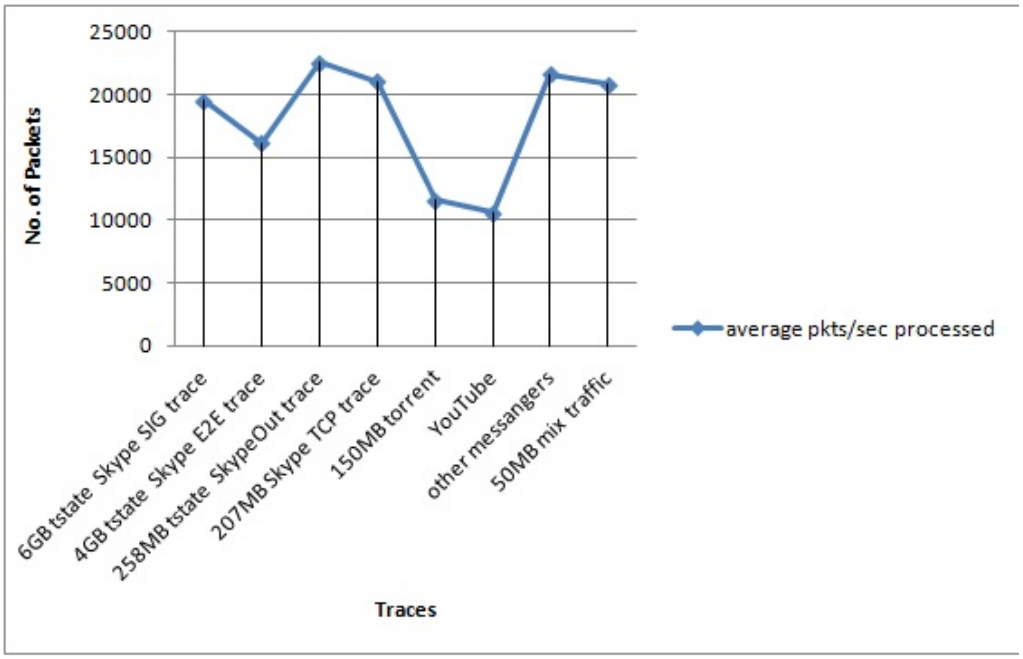

Figure 4.5: Average packets/sec processed by the system on different datasets

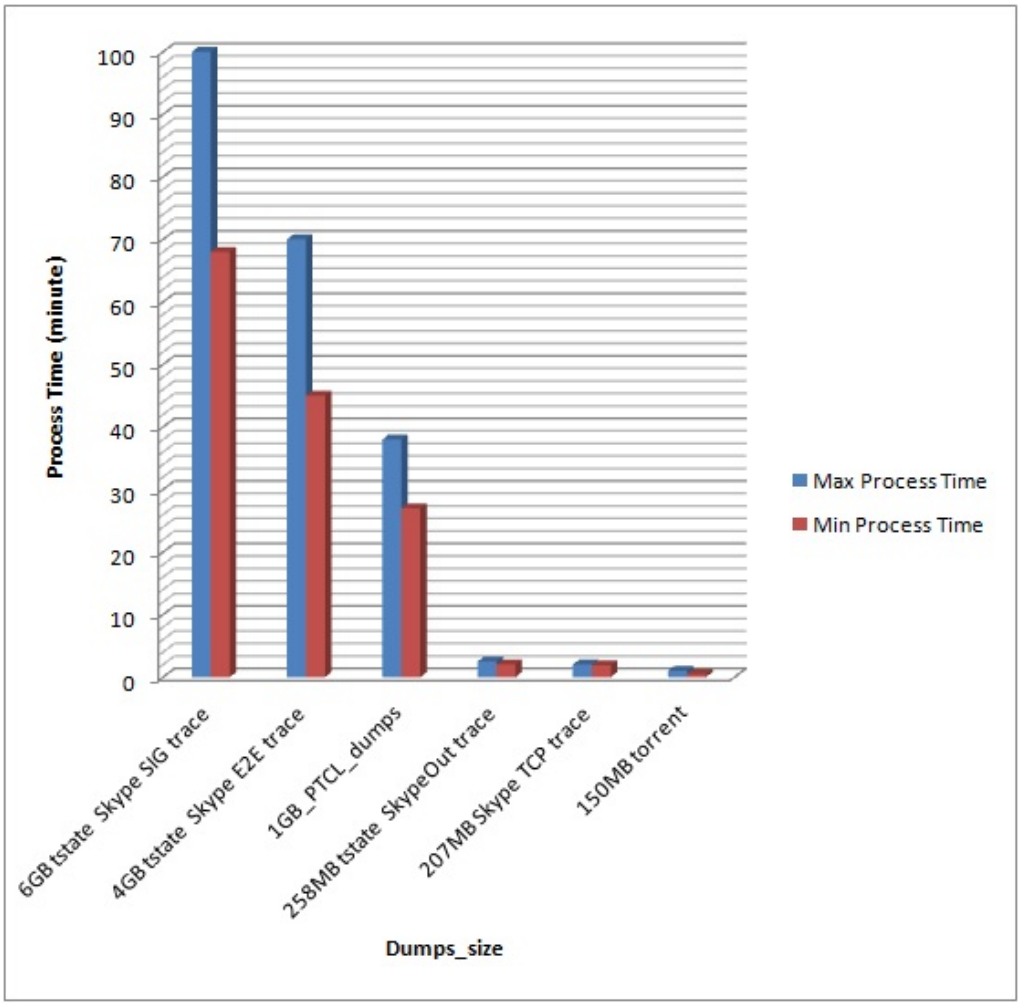

Figure 4.6: Processing time on different datasets

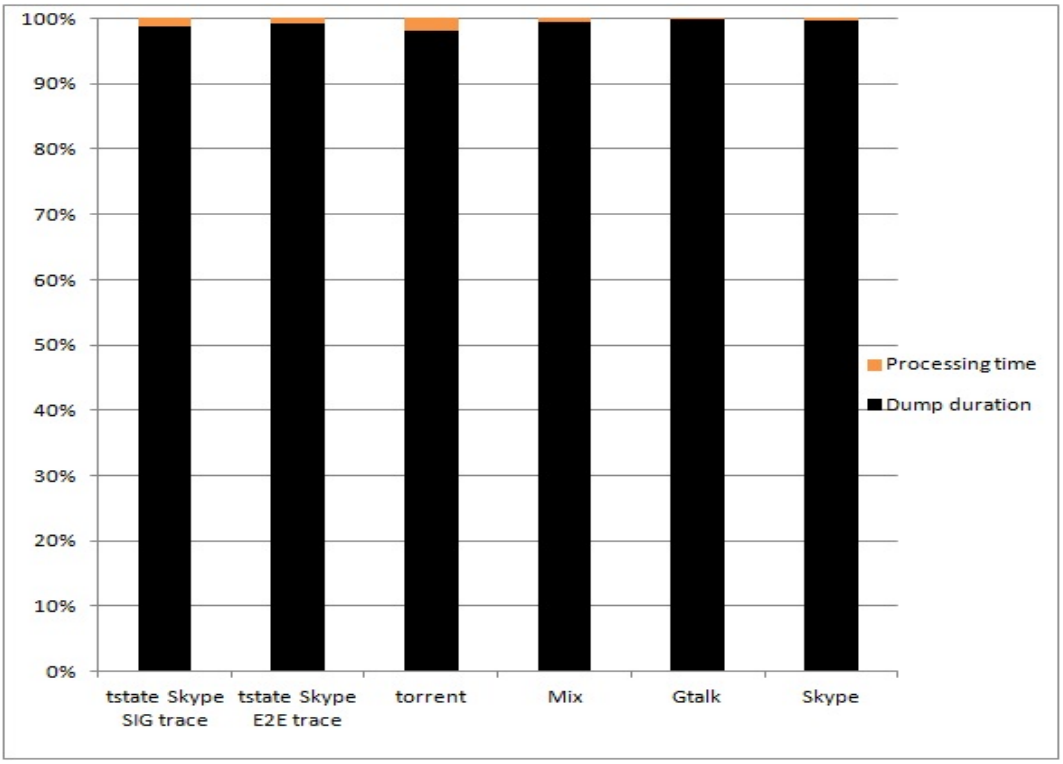

Figure 4.7: Trace time duration and processing time comparison on different datasets

shown in figure 4.6. Large datasets take more time to be processed. Moreover the datasets those have more flows within short periods need more time to be processed. PTCL dumps comparatively have higher processing time as they have more number of flows even in 1GB dumps within few seconds. So for real time implementation at telecommunication authorities gateway we have to implement our system efficiently with parallel programming. Moreover we need powerful servers to handle huge amount of incoming traffic within a few milliseconds. Figure 4.7 provides the comparison between actual duration of the trace and the time taken to process the trace. The duration is quite higher as compare to the processing time of the trace in case of tstat Skype traces and the traces captured from end users computers.

So our system works efficiently and detects VoIP media (voice) flows within 5 seconds from start of the voice conversation. We just need a powerful server and parallel programming to handle huge amount of incoming traffic at telecommunication gateways such as PTA, PTCL gateways.

4.4 Scalability

Scalability is defined and discussed in terms of system usage. We have observed that the system performance is better with respect to accuracy and efficiency. Moreover our system is generic that can detect VoIP traffic regardless of the VoIP application, protocol, codec used, and security mechanism. The system can detect tunneled VoIP such as SSL/TLS and IPSec VoIP. The system can be implemented at one-way or two way network interface. Our system is only specific to VoIP detection so it has better results than other P2P traffic classifiers. So our system is scalable and practically implementable at telecommunication authorities or ISPs gateway with powerful servers and optimized and efficient programming implementation for real time VoIP calls detection. It meets the need of telecommunication authorities for detecting VoIP flows to either prioritize or block traffic from these flows. So we can say, our system is scalable in terms of usage and implementable at any location or site.

4.5 Comparison with existing techniques

Here in this section we compare our technique with the existing VoIP detection techniques. We provide comparison in terms of accuracy and salient features of the technique. We compare our technique with host and flow behavior analysis (HFBA) technique [2], threshold-based detection [3], IPSec tunneled VoIP detection [5], and flow level behavior (FLB) technique [27] in terms of TP, FP, FN. HFBA [2], threshold-based detection [3], and FLB [27] techniques analyzed,tested, and gave more importance to Skype voice traffic. So we chose them for comparing with our techniques on Skype traces. Our technique is threshold-based statistical analysis technique which can detect IPSec tunneled voice calls. So we chose and implement a threshold-based VoIP detection technique presented in [3] and IPSec VoIP detection technique[5] for comparing on all offline captured traces. The IPSec VoIP detection technique has similar features as our technique i.e generic, supported both one-way or two-way interfaces, specific to VoIP detection, can detect IPSec tunneled voice.

Table 4.6 shows the results and comparison on tstat 4GB Skype trace and we have noticed that our system has better results. Table 4.7 shows overall accuracy results of different techniques on offline (table 4.2, table 4.4) and on our VoIP setup traces (table 4.3). Our technique has better results than all other techniques. Our VoIP setup uses G711 codec scheme

for voice encoding. The size of voice packets exceeds from the rang handled by threshold-based VoIP detection technique [3] and IPSec VoIP detection technique [5]. So both of these techniques have 0% TP on our VoIP setup traces. IPSec VoIP detection technique has 25% FP on all offline captures traces which is quite higher and unbearable. Similarly TP of the threshold-based detection technique [3] is quite lower.

Table 4.6: Comparison between our technique and existing techniques w.r.t accuracy on 4 GB tstat Skype traces

Technique	**TP** %	**FN** %
Host and flow behavior analysis (HFBA) [2]	90.28	9.72
Threshold-based detection [3]	79.2	20.8
Flow level behavior (FLB)[27]	55.6	44.4
Our technique	98.86	1.14

We have also compared our technique with existing techniques by features in table 4.8. Our technique has silent features i.e. generic, support both type of interfaces, can detect IPSec tunneled VoIP, and specific to VoIP detection. Moreover we tested and analyzed more VoIP applications than all other techniques. Only the IPSec VoIP detection technique [5] and PDF-PSD VoIP detection technique [8] have similar features. PDF-PSD technique only analyzed and tested Skype voice traffic using 3 codec schemes. IPSec VoIP detection technique [5] do not have good accuracy results as it only depends on packet size. Moreover it only analyzed and tested its own VoIP setup, no other VoIP application is analyzed.

Table 4.7: Comparison between our technique and existing techniques w.r.t accuracy on different captured traces

Technique	**Offline traffic**		**VoIP setup traces**	
	TP %	**FP** %	**TP** %	**FN** %
Threshold based detection [3]	35.7	.0001	0	100
IPSec VoIP detection [5]	73	25	0	100
Our technique	96.56	.00013	100	0

Table 4.8: Overall techniques comparison

Technique	**Year**	**Basic parameters**	**Generic (w.r.t VoIP application or protocol)?**	**Supported interface (one-way, two-way)**	**IPSec VoIP detection?**	**Specific to VoIP detection?**	**VoIP applications tested**
HFBA[2]	2010	No.of ports, packet time	Yes	Two-way	No	Yes	Skype
PDF-PSD [8]	2010	Packet size, packet time	Yes	Both	Yes	Yes	Skype
IPSec VoIP detection [5]	2010	Packet size	Yes	Both	Yes	Yes	Own VoIP setup
C4.5, AdaBoost, SBB-GB [15]	2010	Packet size, time	No	Two-way	Yes	Yes	Gtalk, Skype
PSD-PA [9]	2009	Packet size, ports	Yes	Two-way	No	No	Skype, MSN
K-means classifier [6]	2009	Packet size, time, direction	No	Two-way	No	No	Nil
Statistical thresholds [3]	2008	Packet size, exchange rate	Yes	Both	No	Yes	Skype, MSN, Yahoo, Gtalk
VoIP hidden in web traffic [4]	2008	Request and respond size, time, no. of requests	No	Two-way	No	Yes	Skype, Gtalk
J48, REP tree [12]	2007	Packet size, time, flow duration	Yes	Both	No	No	MSN, Skype
FLB [27]	2006	Packet size, time	No. (specific to some VoIP applications)	Both	No	Yes	Skype, Netmeeting. SIPSoftphone, Kaza
Our technique	2012	Packet size, time	Yes	Both	Yes	Yes	Skype, MSN, Yahoo, Gtalk, Gmail, Zfone, X-lite, Asterisk, Blink, Eyebeam

4.6 Summary

This chapter has provided details of datasets tested, performance of our proposed technique and detailed comparisons between our technique and all other existing techniques. We have observed the results and features of our technique, so after detailed comparisons we can say that the proposed system is better choice for telecommunication authorities and ISPs for VoIP detection.

Chapter 5

Conclusions

5.1 Conclusion

In this thesis, we have proposed practical solution for telecommunication authorities and ISPs to detect VoIP media flows. A generic, robust, efficient, and practically implementable statistical analysis-based solution to detect encrypted, non-encrypted, and any kind of tunneled VoIP media flows using threshold values of flow statistical parameters by giving one-way traffic is proposed in this thesis.

In this thesis, we have provided thorough study of existing VoIP detection techniques especially the statistical analysis-based techniques. We critically analyzed existing work done in VoIP detection field. The limitations are identified in the existing solutions so that we can remove them by our ssytem. Detailed statistical analysis and experiments are done on VoIP applications to remove the limitation in existing solutions and to find an efficient solution for VoIP detection. Different VoIP applications such as Gtalk, Skype, MSN, Yahoo, Zfone, Asterisk etc. are used for VoIP media traffic analysis. PTA, PTCL dumps are also used for analyzing RTP traffic. Different VoIP protocols such as RTP, ZRTP, and SRTP are also examined. We used Wireshark and C programming language with Winpcap library for traffic analysis. The signatures of encrypted tunnels such as SSL and IPSec are also identified for tunneled VoIP analysis. On the basis of these analysis, the statistical parameters are suggested to distinct voice flows from other flows. The threshold values for these statistical parameters are identified for VoIP media (voice) flows detection. Finally a technique is proposed on the basis of these statistical analysis that contains different rules to be true for VoIP flow. Two main algorithms are proposed; one for detecting encrypted, non-encrypted VoIP,

or transport layer tunneled VoIP such as SSL or SSH VoIP and other for detecting voice flows hidden in IP layer tunnels such as IPSec. At the end, the proposed solution is tested on tstat Skype traces, other VoIP application traces such as Gtalk, MSN, Skype, Yahoo, Zfone, X-lite, Eyebeam, and Blink with Asterisk as VoIP server. Own VoIP setup is also established for testing and analysis VoIP by using Asterisk as server and Eyebeam and Blink as clients that communicate both encrypted (SRTP) and non-encrypted voice (RTP). Non-VoIP applications traces i.e. web mails, YouTube, torrents, online games, online TVs, songs, watching movies etc. are also tested. The proposed solution is tested on these traces for accuracy and efficiency.

At the end, the comparison is made between our technique and existing techniques on the basis of accuracy and silent features. The comparisons and results show that our technique is best among all the existing techniques. Our technique has 97.54% TP and .00015% FP. Our technique is generic, robust, efficient and practically implementable that detect the encrypted, non-encrypted and tunneled voice flows from random one-way or two-way traffic. So it is the best choice for telecommunication authorities and ISPs for VoIP detection.

5.2 Future work

In this thesis, we considered the tunnels that are only used for voice communication. We did not consider the tunnels that are used for multi purposes. Our proposed solution does not detect those tunnels that are used for VoIP as well as for other types of traffic simultaneously. We are planning to detect such multipurpose tunnels that also communicate voice traffic. Moreover we are also planning to detect the VoIP hidden in such tunnel.

Bibliography

[1] Bing Li, Shigang Jin, and Moade Ma, "*VoIP Traffic Identification Based on Host and Flow Behavior Analysis*," Journal of Network and Systems Management, Volume 19, 2010.

[2] Bing Li, Shigang Jin, and Moade Ma, "*VoIP Traffic Identification Based on Host and Flow Behavior Analysis*," 2010 internationa conference on Wireless Communication Networking and Mobile Computing (WICOM), Chengdu, China, pp 1-4, 23-25, 2010.

[3] Fauzia Idrees and Uzma Aslam Khan,"*A Generic Technique for Voice over Internet Protocol (VoIP) Traffic Detection*," IJCSNS International Journal of Computer Science and Network Security, VOL.8 No.2, pp 52-59, 2008.

[4] Emanuel P. Freire, Artur Ziviani and Ronaldo M. Salles, "*Detecting VoIP Calls Hidden in Web Traffic," IEEE transaction on network and service management*, Vol no. 5, pp- 210-214, 2008.

[5] Taner Yildirim and Dr. PJ Radcliffe, "*VoIP Traffic Classification in IPSec Tunnels*," 2010 International Conference on Electronics and Information Engineering (ICEIE), Koyoto, Japan, vol. 1, pp VI-151-VI-157, 2010.

[6] Gianluca Maiolini, Giacomo Molina, Andrea Baiocchi, and Antonello Rizzi, "*On the fly Application Flows Identification by exploiting K-Means based classifiers*," Journal of Information assurance and Security, vol 4, pp 142-150, 2009.

[7] Thuy T.T. Nguyen and Grenville Armitage, "*Clustering to Assist Supervised Machine Learning for Real-Time IP Traffic Classification*," IEEE International Conference on Communication, Beijing, China, pp 5857-5862, 2008.

[8] Taner Yildirim and Dr. PJ Radcliffe, "*A Framework for Tunneled Traffic Analysis*," 12th International Conference on Advance Communication Technology (ICACT), Phoenix Park, Korea, vol 2, pp 1029-1034, 2010.

[9] Ying-Dar Lin, Chun-NanLu, Yuan-ChengLai, Wei-HaoPeng, and Po-ChingLin, "*Application classification using packet size distribution and port association*," ELSEVIER: Journal of Networ and Computer Applications, Vol 32, pp 1023-1030, 2009.

[10] M. Dusi, M. Crotti, F. Gringoli, and L. Salgarelli, "*Tunnel Hunter: Detecting application-layer tunnels with statistical fingerprinting*," ELSEVIER: Computer Networks, vol. 53, pp 81-97, 2009.

[11] Riyad Alshammari and Nur Zincir-Heywood, "*Can encrypted traffic be identified without port numbers, IP addresses and payload inspection?*," ELSEVIER: Computer Networks, vol. 55, pp 1326-1350, 2011.

[12] Li Jun, Zhang Shunyi, Liu Shidong, and Xuan Ye, "*Active P2P Traffic Identification Technique*," 2007 International Conference on Computational Intelligence and Security, Harbin, China, pp 37-41, 2007.

[13] Dario Rossi, Marco Mellia, and Michela Meo, "*A Detailed Measurement of Skype Network Traffic*," Proceedings of the 7th international conference on Peer-to-peer systems,USENIX Association, 2008.

[14] Dario Bonfiglio, Marco Mellia, Michela Meo, Nicol'o Ritacca and Dario Rossi, "*Tracking Down Skype Traffic*," INFOCOM 2008, 27th IEEE International Conference on Computer Communication, Pheonix, Arizana, USA, pp 261-265, 2008.

[15] Riyad Alshammari and A. Nur Zincir-Heywood, "*An Investigation on the Identification of VoIP traffic: Case Study on Gtalk and Skype*," 2010 International Conference on Network and Service Management (CNSM), Niagara Falls, Canada, pp 310-313, 2010.

[16] WIRESHARK SampleCaptures available at "http://wiki.wireshark.org/SampleCaptures", accessed on December 2011.

[17] Skype Testbed Traces available at "http://tstat.tlc.polito.it/traces-skype.shtml", accessed on March 2011.

[18] Chun-Ming Leung and Yuen-Yan Chan,"*Network Forensic on Encrypted Peer to-Peer VoIP Traffics and The Detection, Blocking, and Prioritization of Skype Traffics*," 16th IEEE International Workshops on Enabling Technologies: Infrastructures for Collaborative Enterprises, Paris, France, pp 401-408, 2007.

[19] Salman A. Baset and Henning Schulzrinne, "*An Analysis of the Skype Peer-to-Peer Internet Telephony Protocol*," INFOCOM 2006, 25th International Conference on Computer Communication, Bercelona, Spain, pp 1-11, 2006.

[20] P. Renals and G. A. Jacoby, "*Blocking Skype through deep packet inspection*," the 42nd Annual Hawaii International Conference on System Sciencs, HICSS'09, Big Island, HI, pp. 1-5, 2009.

[21] Zhou-li Xu, Zhi-hong Jiang, Song-hai Mo, and Peng-yi Fan,"*Identification of P2P streaming traffic using application signatures*," Application Research of Computers, vol. 26, pp. 2214-2216, 2009.

[22] R. C. Dodge, "*Skype fingerprint*," 41st Hawaii International Conference on System Sciencs, Waikoloa, HI, pp. 399-404, 2008.

[23] Feng LU, Xiao-Lei LIU, and Zhi-Nan MA, "*Research on the characteristics and blocking realization of Skype protocol*," 2010 IEEE International Conference on Electrical and Control Engineering (ICECE), Wuhan, China, pp 2964-2967, 2010.

[24] Robert Birke, Marco Mellia, Michele Petracca, and Dario Rossi, "*Experiences of VoIP Traffic Monitoring in a Commercial ISP*," International Journal of Network Management, vol. 20, pp 339-359, 2010.

[25] D. Adami, C. Callegari, S. Giordano, M. Pagano, and T. Pepe, "*A Real-Time Algorithm for Skype Traffic Detection and Classification*," 9th International Conference on Smart Spaces and Next Generation Wired/Wireless Networking and 2nd Conference on Smart Spaces, Postersburg, Russia, 2009.

[26] M. Peranyi, S. Molnar, "*Enhanced Skype traffic identification, in: ValueTools*," Proceedings of the 2nd International Conference on Performance Evaluation Methodologies and Tools, ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), ICST, Brussels, Belgium,Belgium, pp. 1-9, 2007.

[27] T. Okabe, T. Kitamura, and T. Shizuno, "*Statistical Traffic Identification Method Based on Flow-Level Behavior for Fair VoIP service*", Proceedings of IEEE Workshop on VoIP Management and Security, pp. 35-40, 2006.

[28] Chen-Chi Wu, Kuan-Ta Chen, Yu-Chun Chang, and Chin-Laung Lei, "*Detecting VoIP Traffic Based on Human Conversation Patterns,"* Principles, Systems and Applications of IP Telecommunications, Services and Security for Next Generation Networks, Second International Conference, IPTComm 2008, Heidelberg, Germany, 2008.

[29] Yuanchao Lu and Ye Zhu, "*Correlation-Based Traffic Analysis on Encrypted VoIP Traffic*," Second International Conference on Networks Security, Wireless Communications and Trusted Computing (NSWCTC), Wuhan, Hubei, China, Vol. 1, pp 45-48, 2010.

[30] Hyunchul Kim, kc claffy, Marina Fomenkov, Dhiman Barman, Michalis Faloutsos, and KiYoung Lee, "*Internet Traffic Classification Demystified: Myths, Caveats, and the Best Practices*," ACM CONEXT 2008, Madrid, Spain, 2008.

[31] Olivier Verscheure, Michail Vlachos, Aris Anagnostopoulos, Pascal Frossard, Eric Bouillet, and Philip S. Yu, "*Finding Who Is Talking to Whom in VoIP Networks via Progressive Stream Clustering*," 6th International Conference on Data Mining (ICDM), Hong Kong, China, pp 667-677, 2006.

[32] Tuneesh Lella and Riccardo Bettati, "*Privacy of Encrypted Voice-over-IP*," IEEE International Conference on System, Man and Cybernetics (ISIC), Quebec, Canada, pp 3063-3068, 2007.

[33] Charles V. Wright, Scott E. Coull, and Fabian Monrose,"*Traffic Morphing: An Efficient Defense Against Statistical Traffic Analysis*," 16th IEEE Networks and Distributed Security Syposium (NDSS), pp 237-250, 2010.

[34] Ye Zhu and Huirong Fu, "*Traffic analysis attacks on Skype VoIP calls*," ELSEVIER: Computer Communications, vol. 34, pp 1202-1212, 2011.

[35] Chang-Yong Lee, Hwan-Kuk Kim, Kyoung-Hee Ko, Jeong-Wook Kim, and Hyun Cheol Jeong, "*A VoIP Traffic Monitoring System based on NetFlow v9*," International Journal of Advance Science and Technology, vol 4, 2009.

[36] R. Muralishankar, R. Venkatesha Prasad, Vijay S, and H. N. Shankar, "*Order Statistics for Voice Activity Detection in VoIP*," 2010 IEEE International Conference on Communication (ICC), Cape Town, South Africa, pp 1-6, 2010.

[37] Jian Feng, "*Research on the Technology of Peer-to-Peer Traffic Classification*," 2010 International Symposium on Computer, Communication, Control and Automation (3CA), Tainam, Taiwan, vol. 1, pp 491-494, 2010.

[38] Roberto Barbieri, Danilo Bruschi, and Emilia Rosti, "*Voice over IPsec: Analysis and Solutions*," Proceding of the 18th annual Computer Security Application Confrence (ACSAS), Las Vegas, pp 261-270, 2002.

[39] Luca Deri, "*Open Source VoIP Traffic Monitoring*," System Administration and Networking Engineering (SANE), Nether Lands, 2006.

[40] E. P. Freire, A. Ziviani, and R. M. Salles, "*Detecting Skype flows in Web traffic*," NOMS 2008, Proceedings of the 2008 IEEE/IFIP Network Operations and Management Symposium, 2008.

[41] Federal Communications Commission, "*Voice over Internet Protocol*," http://www.fcc.gov/voip/

[42] Jun-peng Mao, Yan-li Cui, Xiang-jie Ma and Yan-feng YU, "*Traffic measurement and characteristics finding of Skype network*," Computer Engineering, vol. 34, pp. 142-144, 2008.

[43] M. Baugher, D. McGrew, M. Naslund, E. Carrara, and K. Norrman, "*The secure real-time transport protocol (srtp)*", United States, 2004.

[44] T. Karagiannis, A. Broido, M. Faloutsos, and K. claffy, "*Transport layer identification of P2P traffic*," IMC '04, Proc. 4th ACM SIGCOMM Conference on Internet Measurement, pp. 121-134, 2004.

[45] S. Guha, N. Daswani, and R. Jain, "*An experimental study of the Skype peer-to peer VoIP system*," IPTPS'06, Proc. 5th International Workshop on Peer-to-Peer Systems, pp. 1-6, 2006.