A Framework for Secure VoIP

Mandeep Singh
Research Scholar
Department of Computer Science & Engineering
Ganga Institute of Technology & Management, Kablana

Neetu Sharma
Head of the Department
Department of Computer Science & Engineering
Ganga Institute of Technology & Management, Kablana

Abstract

VoIP stands for Voice over Internet Protocol and is a way to carry voice traffic over computer networks like the Internet. Over the last decade VoIP has become increasingly popular, gaining millions of subscribers every year (e.g. LINE and WECHAT provide voicecall facilities) and has certainly caught the eye of telecommunication service providers all over the world. The driving factor for the success of VoIP is cost reduction, both for users and providers. But VoIP doesn't only bring reduced costs it also brings threats and vulnerabilities since it is IP based it's susceptible to large number of threats. The threats include spoofing or identity theft and call redirection, making data integrity a major risk. Therefore authentication and encryption techniques can be used to design a framework which can survive the possible threats. In this security framework authentication is implemented first to authenticate the true user and then cryptography techniques is used to safely transmit the information stream over the network. The authentication part will be implemented using biometrics because it is not possible to theft anyone's physical features.

Keywords: VoIP, PSTN, SIP

I. INTRODUCTION

VoIP has fast emerged as a standard for voice communication using the Internet. As VoIP uses the existing IP network, it dramatically reduces cost of communication typically with traditional PSTN. In addition, ease of deployment and reduced communication hardware make VoIP a compelling solution for voice communication on the Internet. Further, VoIP provides a flexibility of value-added and personalized services for defining customized solutions. As a result, most of the control which existed in PSTN's central infrastructure has been transferred to the end devices by deploying the VoIP communication infrastructure. With the advent of VoIP technology, an increasing number of telecommunication service providers have stated to integrate VoIP solutions into their systems and provide VoIP services to their customer base. Equipment manufacturers and end users have greatly benefited from performance advancements, cost reduction, and feature support provided by the VoIP technology.

VoIP is a technology for transmitting voice packets on the existing IP network. Unlike PSTN, an IP network is packet switched. In PSTN, when a phone call between two parties is initiated, there exists a physical circuit connecting the two parties. After the call is established, the parties communicate and the circuit is reserved until the parties finish the communication. In contrast, on an IP network, all communication is carried out using IP packets. When a calling party communicates with a called party, the analog signals are digitized, encoded, and packed into an IP packet at the transmitting end and converted back to analog signals at the receiving end.

VoIP is adding a third dimension to voice communication with the PSTN and cellular networks being the other two. A call can be made to any PSTN phone and mobile phone anywhere in the world using VoIP. Although certain services can only function on computer or a special VoIP phone; others allow a caller to use a traditional phone with an adapter. VoIP promises to enable migration of the existing circuit-switched, public switching telecom network to a packet-switched network. With VoIP, widespread acceptance by telecommunication markets of all sizes, advanced features have started emerging. However, the convergence of the voice and data worlds introduces not just opportunities but also security risks. The much lower cost and greater flexibility are key factors luring enterprises to transition to VoIP. VoIP should not, however, be installed without careful consideration of the security problems it can introduce.

Security issues in VoIP are unique and, in most cases, quit complex. This article aims to provide an overview of VoIP security issues including basic VoIP architecture, existing defense mechanisms, and current attacks, as well as an outlook on potential attacks such as SPIT and their possible solutions.

We may consider a VoIP call as a three−phase process:

- Establishment
- Conversation
- Termination

The first and the third phases typically make use of a signalling protocol, such as the Session Initiation Protocol (SIP), while the second utilizes the Real−Time Protocol (RTP) to transport the media data. Therefore, this project handles the signalling protocol security and the data transport protocol security independently. The main goal of this document is the description of a suitable solution to achieve this security without effecting the performance of the model.

Mobility aspects of the model are not explicitly considered in this project, although this paper may be one of the bases for future work in this area, since the proposed solution is based on the use of mobile devices operating in wireless environments.

Regarding practical work and tests of the model, this project in mainly focused on the security for the media stream (by using the Secure Real−Time Protocol, SRTP). However, thorough theoretical work has also been performed, which includes other aspects as said above, such as the establishment and termination of the call (using the Session Initiation Protocol, SIP) and the key−management protocol to be used.

A. Security Attacks

A security attack is defined as an assault on system security that derives from an intelligent threat (which might exploit a vulnerability), and compromises the security of information owned by a organization. Security Attacks are divided into two main groups: Passive Attacks and Active Attacks.

B. Passive Attacks

Those whose goal is to obtain information that is being transmitted. Passive Attacks are divided into two main groups:

- Release of message contents: Interception of the content (possibly sensitive) of a message.
- Traffic Analysis: Interception for observing the patterns of the messages to guess the nature of a communication.

C. Active Attacks:

Those which involve some modification or alteration of the data stream, or the creation of a false stream. Active Attacks are in turn divided into four groups:

- Masquerade: It implies one entity pretending to be a different entity.
- Replay Attack: It consists of the capture of sensitive data, and its subsequent retransmission to produce an unauthorized effect.
- Modification of messages: It implies the alteration, deletion, delay, or reordering of some portion of a message, producing an unauthorized effect.
- Denial of Service (DoS) Attack: DoS attack prevents or inhibits the normal use or management of communication facilities by disabling or overloading them.

Passive Attacks are difficult to detect, since they do not imply alteration of the data. Thus, the solution is the prevention of these attacks, and the mechanism used is encryption.

On the other hand, Active Attacks are difficult to prevent, since that would imply the physical protection of resources and paths. Therefore, the solution is to detect and recover from these attacks.

D. Processes Involved

Authentication Process: This section describes the authentication process by means of the establishment of a call. The mentioned process outlines the situation where the goals of the implementation are already met. Systems that communicate try to initiate a call on the request of one of the EPs that are registered to it. The communication between the two devices is part of the call signaling to establish a call.

E. Biometric Cryptosystem:

Cryptography provides the secure manner of information transmission over the insecure channel. It authenticates messages based on the key but not on the user. It requires a lengthy key to encrypt and decrypt the sending and receiving the messages respectively. But these keys can be guessed or cracked. Moreover, Maintaining and sharing lengthy, random keys in enciphering and deciphering process is the critical problem in the cryptography system. The above mentioned problem is solved by a Biometric cryptosystems. Biometric cryptosystems combine cryptography and biometrics to benefit from the strengths of both fields. In such systems, while cryptography provides high and adjustable security levels, biometrics brings in non-repudiation and eliminates the need to remember passwords or to carry tokens etc. In biometric cryptosystems, a cryptographic key is generated from the biometric template of a user stored in the database in such a way that the key cannot be revealed without a successful biometric authentication.

II. REVIEW OF LITERATURE

In Year 2006, Matthew Simon proposed “Forensic Computing Implications of Voice over IP”. The Author concluded that imaging memory for forensic purposes is gaining recognition asa an area in need of further research. The scope of memory forensics is large because it has potential to add a diverse source of potential evidence.

In Year 2009, Hesham N. Elmahdy performed a work,”The impact of Packet Size and Packet Dropping Probability on Bit Loss of VoIP Networks”. The Author introduced the effect of red packet dropping probability and the packet size on the Fairness index and lost packets via a computer simulation and concluded that SD is conversely proportional to the Fairness index. The dropping probability of the red packets has a great effect on the packet loss.

In Year 2012, S. Alshomrani and his teammates performed a work," QoS of VoIP over WiMAX Access Networks". The Author evaluated the performance of VoIP over the WiMAX networks. Different parameters such as jitter, MOS value, packet end-to-end delays and packets sent and received were used to measure the performance of VoIP over WiMAX. The research showed that VoIP applications can perform better under the exponential traffic distribution.

In Year 2012, Ge Zhang performed a work "Unwanted Traffic and Information Disclosure in VoIP Networks". The Author concluded that:

- Achieving security and privacy for VoIP is more difficult than for traditional PSTN.
- Unbalanced resource consumption between client side and server side leads to DoS attacks.
- It is more difficult to classify voice spam than text spam.
- Side channels in VoIP traffic disclose secret information.
- Security products for VoIP should be designed taking efficiency into account.

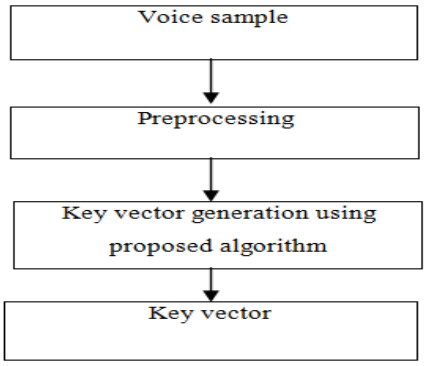
In Year 2014, Weiwei Zhang and his teammates performed a work "Study on the QoE for VoIP Networks". The author constructed a VoIP network experiment platform to achieve the real network delay, jitter and packet loss characteristic change by adjusting the parameters of network simulation software. There are three main factors that affect quality is speech codec network packet loss and network delay. He compared the QoE value of the voice quality for three codec under different packet loss rate and different delay.

III. PROBLEM DEFINITION

VoIP is fastly gaining popularity as a standard for communication using internet. As VoIP uses the existing IP network, it dramatically reduces cost of communication typically with traditional PSTN but also as it uses internet protocol ,it welcomes different types of security threats to the user. Therefore a framework should be designed so that enhances the quality of service and minimizes the security threats. In this work two processes are involved to design the framework. The processes are Authentication process and Biometric Cryptosystem. Authentication Process describes the authentication process by means of the establishment of a call. Cryptography provides the secure manner of information transmission over the insecure channel. It authenticates messages based on the key but not on the user. It requires a lengthy key to encrypt and decrypt the sending and receiving the messages respectively. But these keys can be guessed or cracked. Moreover, Maintaining and sharing lengthy, random keys in enciphering and deciphering process is the critical problem in the cryptography system. The above mentioned problem is solved by a Biometric cryptosystems. Biometric cryptosystems combine cryptography and biometrics to benefit from the strengths of both fields. In such systems, while cryptography provides high and adjustable security levels, biometrics brings in non-repudiation and eliminates the need to remember passwords or to carry tokens etc. In biometric cryptosystems, a cryptographic key is generated from the biometric template of a user stored in the database in such a way that the key cannot be revealed without a successful biometric authentication.

IV. FLOW CHART USED

Cryptography provides high and adjustable security levels, biometrics brings in non-repudiation and eliminates the need to remember passwords or to carry tokens etc. In biometric cryptosystems, a cryptographic key is generated from the biometric template of a user stored in the database in such a way that the key cannot be revealed without a successful biometric authentication.



Key generation using voice samples

A human can easily recognize a familiar voice. However, getting a computer to distinguish a particular voice among others is a more tricky task. Immediately, several problems arise when trying to write a voice recognition algorithm. The majority of these difficulties are due to the fact that it is almost impossible to say a word exactly the same way on two different occasions. Some factors that continuously change in human speech are how fast the word is spoken, emphasizing different parts of the word, etc. Furthermore, suppose that a word could in fact be said the same way on different occasions, and then we would still be left with

another major dilemma. Namely, in order to analyze two sound files in time domain, the recordings would have to be aligned just right so that both recordings would begin at precisely the same moment. We start comparison by storing two voices in .wav files. Then we plot both signal and try to match them. Directly comparison done here by producing wav files through sampling which is done by calculating it's Fourier transform. Next we plot it's power spectra and then truncate it to form a new power spectra with differences like noise and height of peaks which is to be normalized resulting in a new power spectra. Using mathematical functions we compute and plot an average power spectrum which is also normalized to compare it with two individual voices giving us desired results.

The voice comparison and analysis in our proposed system has been done in the following way illustrated with an example:

We have recorded the word ``hello" twice by the same speaker. The sounds are stored in ``test1.wav" and ``test2.wav". We start by reading them into MATLAB vectors Y1 and Y2.

```
Y1=wavread('test1');          Y2=wavread('test2');
```

Next, we plot both signals in one figure. To do this we type

```
subplot(2,1,1);          plot(Y1)
subplot(2,1,2);          plot(Y2)
```

The result is shown in below figure:

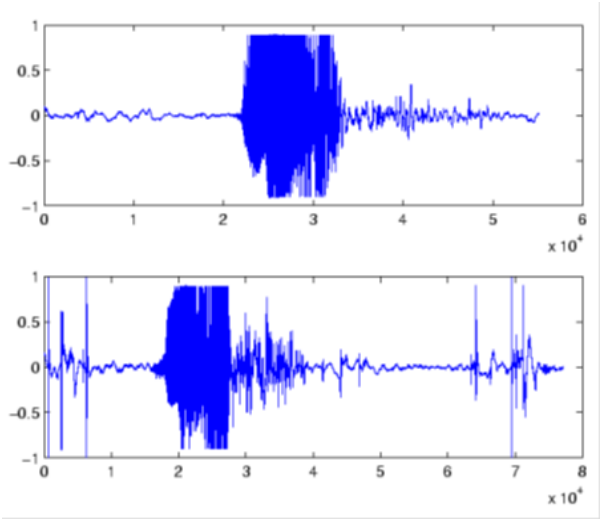
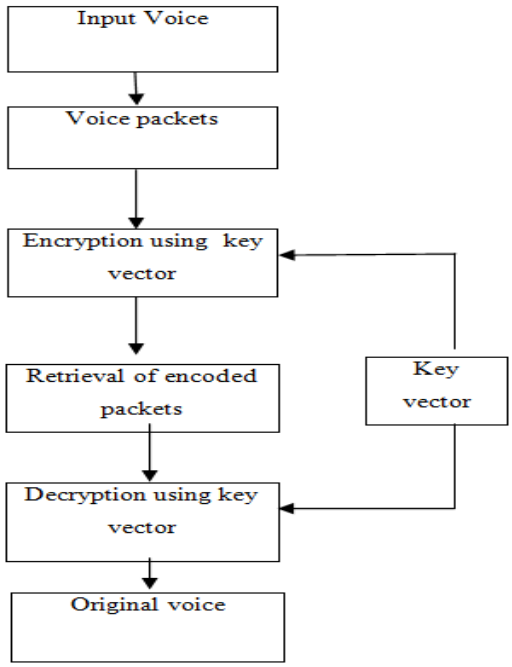


Fig. 1: The two signals Y1 and Y2



Encryption and Decryption using key

A. Key Generation from Voice Samples

In this Section we explain the Key Generation

Algorithm Assumptions

- Kl → length of the AES key
- Mp → voice samples set
- Kl → Key length
- Np → Size of voice samples set
- S → Seed value
- Sl → seed limit.
- m → (x,y) – co-ordinate of a voice sample
- Kv → Key Vector

1) Step 1: The Extracted voice samples are Represented as
- Vs = { vi }i=1,…, Np

2) Step 2: The initial key vector is defined as follows,
- Kv = { xi : p(xi) } i=1, ….., Kl

Where

- p(x) = Vs[I % Np] + Vs[(i+1) % Np]+ S

i=1,…. , Kl

3) Step 3: Initial value of S is equal to total Number of Voice samples. The value of S will be dynamically changed as follows
- S = Kv (i) % Sl , -1 < i< Kl

4) Step 4: Initial key vector (Kv) is converted into a matrix Km of size Kl / 2 * Kl / 2
- Km = (aij) Kl / 2 * Kl / 2

5) Step 5: A intermediate key vector is generated as follows
- KIV = { Ki : (m(ki) } i=1,…. Kl

Where

- m(k) = | Aij | ,
- Aij = Km i,j : i+size,j+size , -1 < i< Kl/2
- Aij is a submatrix formed from the key matrix.

6) Step 6: Final key vector is formed is

Sv = { 1, if KlV [i] > mean(KlV)
,0 otherwise }

V. CONCLUSION

In this paper, a framework was proposed to design a secure framework for safe communication using voice over Internet. The authentication and safety of voice data was ensured by using the voice biometrics for generating the key for encrypting the voice packets. Since physical features of human are unique ,hence this framework minimizes the fraudulent calls and unauthorized access to the voice data.

REFERENCES

[1] B. Goode, "Voice Over Internet Protocol (VOIP)". Proceedings of thee IEEE, VOL. 90, NO. 9, Sept. 2002.
[2] "Voice over Internet Protocol" from http://en.wikipedia.org/wiki/Voice_over_IP
[3] "Security Risk Factors in IP Telephony Based Networks "http://www.syssecurity.com/archive/papers/Security_Risk_Factors_with_IP_Telephony_based_Networks.pdf
[4] Anil K. Jain, Arun Ross, Salil Prabhakar "An Introduction to Biometric Recognition" IEEE Transactions on circuit and systems for video technology, Vol 14, No. 1, January 2004.
[5] N.K. Ratha, J.H. Connell, and R.M. Bolle, "A Biometrics-Based Secure Authentication System," Proc. IEEE Workshop Automatic Identification Advanced Technologies, pp. 70-73, Oct. 1999.
[6] Christina Chalastanis; A work parallel to this one, covering VoIP security problems in enterprise networks; Security of SIP-based Voice over IP in enterprise networks, December 2005
[7] VOIPSA, Inc.; A portal of VoIP Security Alliance - organization carrying out many VoIP security related projects; http://www.voipsa.com, 2006
[8] Biggs, P. (2007). The status of voice over internet protocol (VoIP) worldvide, 2006. ITU.
[9] G. R. Doddington. Speaker recognition—Identifying people by their voices. Proceedings of the IEEE, 73(11):1651–1664, 1985.
[10] Q. Li and A. Tsai. A matched filter approach to endpoint detection for robust speaker verification. In Proceedings of IEEE Workshop on Automatic Identification Advanced Technologies (AutoID'99), pages 35–38, October 1999.
[11] A. E. Rosenberg and F. K. Soong. Recent research in automatic speaker recognition. In Advances in Speech Signal Processing, pages 701–738, New York: Marcel Dekker, 1992.
[12] Auckenthaler, R., Carey, M., and Mason, J. Speaker-centric score normalization and time pattern analysis for continuous speaker verification. In Proceedings of ICASSP 2000, the IEEE International Conference on Acoustics, Speech, and Signal Processing (Istanbul, Turkey, June 5–9). IEEE Press, Piscataway, N.J., 2000, 48–50.
[13] Markowitz, J. Ieri, oggi, domani: Speaker recognition yesterday, today, and tomorrow. In Proceedings of COST250 Workshop on Speacker Recognition in Telephony (Rome, Italy, Nov. 10). European Co-operation in the Field of Scientific and Technical Research, 1999.