ISSN (Print) : 0974-6846
ISSN (Online) : 0974-5645

A Comprehensive Survey of Security Issues and Defense Framework for VoIP Cloud

Ashutosh Satapathy* and L. M. Jenila Livingston

School of Computing Science and Engineering, VIT University, Chennai - 600127, Tamil Nadu, India;
ashutosh.satapathy2013@vit.ac.in, jenila.lm@vit.ac.in

Abstract

Voice over Internet Protocol (VoIP) is an advanced telecommunication technology which transfers the voice/video over high speed network that provides advantages of flexibility, reliability and cost efficient advanced telecommunication features. Still the issues related to security are averting many organizations to accept VoIP cloud environment due to security threats, holes or vulnerabilities. So, the novel secured framework is absolutely necessary to prevent all kind of VoIP security issues. This paper points out the existing VoIP cloud architecture and various security attacks and issues in the existing framework. It also presents the defense mechanisms to prevent the attacks and proposes a new security framework called Intrusion Prevention System (IPS) using video watermarking and extraction technique and Liveness Voice Detection (LVD) technique with biometric features such as face and voice. IPSs updated with new LVD features protect the VoIP services not only from attacks but also from misuses.

Keywords: Defense Mechanisms, Liveness Voice Detection, VoIP Cloud, Voice over Internet Protocol, VoIP Security Issues

1. Introduction

The rapid progress of VoIP over traditional services is led to a situation that is common to many innovations and new technologies such as VoIP cloud and peer to peer services like Skype, Google Hangout etc. VoIP is the technology that supports sending voice (and video) over an Internet protocol-based network[1,2]. This is completely different than the public circuit-switched telephone network. Circuit switching network allocates resources to each individual call and path is permanent throughout the call from start to end. Traditional telephony services are provided by the protocols/components such as SS7, T carriers, Plain Old Telephone Service (POTS), the Public Switch Telephone Network (PSTN), dial up, local loops and anything under International Telecommunication Union. IP networks are based on packet switching and each packet follows different path, has its own header and is forwarded separately by routers. VoIP network can be constructed in various ways by using both proprietary protocols and protocols based on open standards.

1.1 VoIP Layer Architecture

VoIP communication system typically consist of a front end platform (soft-phone, PBX, gateway, call manager), back end platform (server, CPU, storage, memory, network) and intermediate platforms such as VoIP protocols, database, authentication server, web server, operating systems etc. It is mainly divided into five layers as shown in Figure1.

1.2 VoIP Cloud Architecture

VoIP cloud is the framework for delivering telephony services in which resources are retrieved from the cloud data center through web applications and software, instead of a direct link to server[3]. Information and applications are stored on cloud servers in a distributed fashion. Apart from cloud computing characteristics such as on demand service, resource pooling, optimize resource allocation, pay as you go, elasticity and scalability[4,5], VoIP cloud contains mainly six components as shown in Figure 2.

**Author for correspondence*

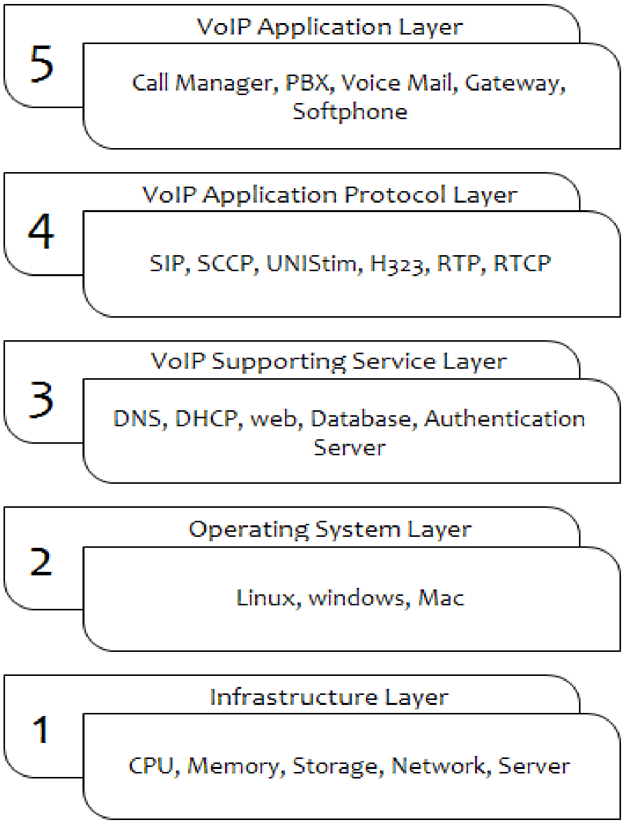
5
VoIP Application Layer
Call Manager, PBX, Voice Mail, Gateway, Softphone
4
VoIP Application Protocol Layer
SIP, SCCP, UNIStim, H323, RTP, RTCP
3
VoIP Supporting Service Layer
DNS, DHCP, web, Database, Authentication Server
2
Operating System Layer
Linux, windows, Mac
1
Infrastructure Layer
CPU, Memory, Storage, Network, Server

Figure 1. VoIP layer architecture.

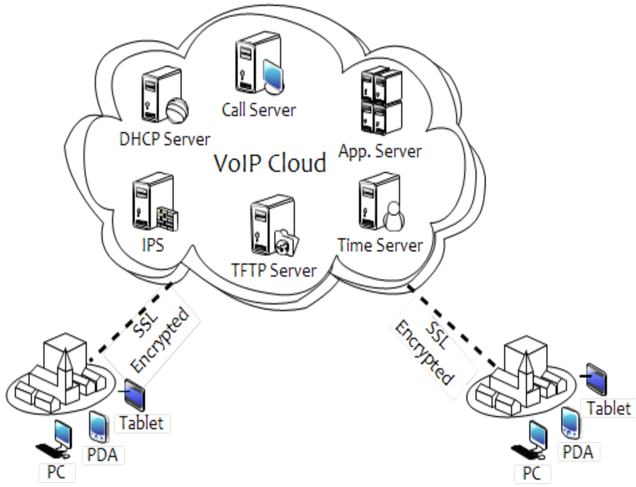

Figure 2. VoIP cloud architecture.

1.2.1 Call Server

Phones are registered with this component. It handles security and admission control while connecting the phones. The Voice data of a call carried by the transport protocol may or may not flow through the call server.

1.2.2 DHCP Server

It is used for dynamically distributing network configuration parameters such as Internet Protocol (IP) address, address of TFTP server etc.

1.2.3 Application Server

These servers are designed to install, host and operate applications and provide services to end users, IT industries and organizations.

1.2.4 Time Server

The main principle of time server is to maintain synchronization over the network. The actual time from server clock is distributed to its clients using a computer network.

1.2.5 TFTP Server

It helps to update the network configuration used by the phones, routers, firewalls and perhaps provide a setting file that might contain operational parameters for VoIP network. e.g., software updates, codec used in a particular region.

1.2.6 Intrusion Prevention System (IPS)

It monitors networks and systems behavior for malicious instances. The major roles of intrusion prevention systems are to find out suspicious instances and their log information, try to block/stop them and report to concern admin.

2. Literature Review

VoIP technology was started in February 1995 by Vocaltec, Inc. in Israel. It transfers the voice over high speed network, cheaper comparing to PSTN and reachable to everywhere through internet by loon developed by Google with 4G LTE speed[6].

2.1 VoIP Security Issues

VoIP transfers the voice over the data network through different network elements such as switches and routers. Connecting PSTN to internet i.e. VoIP as a carrier for voice/video traffic, the security problems are not only common in circuit switch network (PSTN, POTS) such as eavesdropping (tapping) and toll fraud attack but also problems related to IP network. Security issues in VoIP are broadly classified into three categories.

2.1.1 Real Time Issues

From last decade onwards, VoIP is used for several illegal activities such as hacking, terrorism, match fixing etc. Recently in October 2014, phone Hackers had broken into the phone network of the company, Foreman Seeley Fountain Architecture and routed $166, 000 worth of calls from the firm to premium rate telephone numbers in Gambia, Somalia and Maldives. It would have taken 34 years for the firm to run of those charges legitimately, based on its typical phone bill.

2.1.2 Network Related Issues

Attacks related to destroy, block, expose, alter, disable, steal or gain unauthorized access to information in VoIP network (e.g. threats include social, denial of service, service abuse, physical access, interruption of service etc.) are listed in Table 1 followed by different types of attacks[7,8].

2.1.3 Voice Related Issues

As VoIP system carries voice traffic, so victim's voice can be mimicked by an attacker/intruder. A talking and singing robot that mimics human vocalization, developed by M. Kitani, Kagawa University is vulnerable to VoIP communication[9].

2.2 VoIP Attacks

This section deals with different types of VoIP attacks.

2.2.1 Physical Attacks

The attacker performs this attack by stealing, breaking network equipment or direct control over equipment by getting unauthorized access to prohibited area for seeking of information. Some of the physical attacks are dumpster diving, shoulder surfing, hardware key logger and overt access etc. It can be prevented by keeping the documents and records safely inside locker and electronic equipment must be password protected. At last, outer layer security can be provided by deploying security guards at enter and exit points.

2.2.2 MAC Spoofing

The technique of masking a MAC address upon actual MAC address through software emulation is known as MAC spoofing. Here the hacker's system is taken over MAC address of one of the node which is already configured and permitted as VoIP end device by disconnecting or turning off it from rest of the network. It can be prevented by number of ways[10]. When ARP packet arrives, direct extraction of MAC address from LAN card and from OS registry; Compare the MAC address of LAN card with OS. If it doesn't match, then delete the entry from OS registry. Lock down the system by registering its MAC address with a DHCP IP address. At last secure the communication channel by encrypting it.

Table 1. VoIP network threats classification

Threat Type	Description
Social threats	These threats point straight against individuals such as misconfigurations, security holes or defective protocol implementation in VoIP system. (e.g., Phishing, Theft of identity or Service, Social engineering, Spam etc.)
Eavesdropping, interception and modification threats	These threats include illegal/ Un-authorization access and modification of signaling and transport message. (e.g., Call rerouting, interception of RTP sessions etc.)
Denial of service threats	DoS threats repudiate individual access to VoIP services. DDOS attacks strike all of user's or business transmission potentials. (e.g., SYN/UDP floods, ICMP floods, etc.)
Service abuse threats	These threats cause inappropriate utilization of VoIP services when those facilities are provided for business purposes. (e.g., toll fraud and billing avoidance etc.)
Physical access threats	These threats are illegal physical access to VoIP devices or physical layer of the VoIP network. (e.g., Hardware key logger, theft of media, retrieval of discarded stuffs etc.)
Interruption of services threats	These threats cause VoIP services/ facilities to unviable and unavailable. (e.g., power loss due to bad climate, resource consumption due to over purchase/ extra subscription, issues that degenerate call quality etc.)

2.2.3 ARP Spoofing

Hacker spreads forgery Address Resolution Protocol (ARP) packets inside VoIP network by modifying ARP buffer. Here, attacker binds own system MAC address with IP address of genuine server which causes the traffic imply for server is diverted to attacker. It advances hacker not only listen to VoIP calls but also reply and terminate the VoIP calls intended for other. ARP poisoning followed by denial service threats or eavesdropping, interception or modification threats which cause severe damages to victim. So, Enhanced ARP can be implemented to prevent ARP spoofing[11].

2.2.4 IP Spoofing

Attacker gets into the VoIP network by tricking the IP address of any authorized machine which helps him to spread malicious message inside the network. IP spoofing helps attacker to launch further attacks such as DoS attack, theft of services, toll fraud etc. by impersonating authorized host inside VoIP network. Basically IP spoofing can be prevented with maximum probabilities by configuring broader gateway router. First, router disallows incoming packets for destination address coming from source address within one network. Second, router disallows to send packets from local network to another; those don't have source addresses within that local address range. Y. Ma developed an effective trace route based method for counter measure against IP spoofing and it is worked with trusted adjacent nodes information i.e. acceptance of packets for a node is completely depends upon trace route result from its adjacent nodes[12].

2.2.5 ICMP Flood

Internet Control Message Protocol (ICMP) is one of the network layer protocols that carry error and query messages sent by either intermediate nodes or end node. Attacker tries to overflow the receiver cache by flood the respective node with ICMP packets. It forces the node to drop successive ICMP packets until free space available at node's cache even if request packets come from genuine node. Routers are configured to set optimum points for traffic coming from different networks. It will help the routers to not only block unnecessary ICMP packets by matching ICMP requests and responses but also prevent cache overflow. The VoIP system must be configured separate VLAN for packets originating within a single network which are monitored by firewall. Barbhuiya et al. have developed an error detection framework to identify different types of ICMP attack[13]. It consists of two modules. Verification module verifies origination of ICMP packets and Congestion check module extracts bandwidth utilization information using Simple Network Management Protocol (SNMP).

2.2.6 TCP/ UDP Floods

In TCP flooding attack, hacker creates huge number of SYN packets with abnormal source IP addresses and sends to receiver. Receiver node allocates space in its Transmission Control Buffer (TCB) to each SYN requests. In response to SYN packets, receiver sends SYN+ACK packets and waiting for ACK packets. The SYN+ACK packets carry abnormal IP addresses cause failure to receive ACK packets which prevents receiver node to clear TCP SYN requests from buffer and buffer to overflow later. Attacker can use TCP flood attack against VoIP signaling protocol such as H.323 and SIP; as both are connection oriented protocols. Haris et al. have succeed to detect TCP flood attack in communication by analyzing payload and unusable area of the HTTP protocol (e.g., port, flags, source IP, header length)[14].

In UDP flood attack, large number of UDP packets are created with arbitrary source addresses and port numbers and then sends to victim node. Receiver node will check whether any processes are running on those ports and find most of the ports are closed. In reply, receiver node creates large number of destination unreachable packets. Increase the number of ICMP packets causes the victim node and the network to overflow. The UDP flood attack prevents genuine nodes to communicate the victim node at a particular span of time. Attacker can use UDP flood attack against VoIP transport protocol such as RTP and RTCP; as both are connection less protocol. Bardas et al. proposed a proportional packet rate assumption technique to differentiate UDP traffic for detecting forge IP addresses responsible for UDP flood attacks[15].

2.2.7 TCP/ UDP Replay

First, attacker tries to obtain network sensitive information such as session cookies, password, voice data, signaling data. The information captured by sniffing tools can be used by attacker to take over the ongoing session. Sometime victim's voice can be impersonated by directly playing back recorded voice data or slightly modifying voice data and send to destination which helps the hacker to retrieve more information between caller and callee. Encrypt the sessions is the best way to stop penetration. Ali et al. proposed an enhanced port knocking technique to block TCP replay and port scanning attacks[16]. It is worked on source port sequences authentication instead of destination port sequence number.

2.2.8 SIP Registration Hijacking

VoIP phones use SIP or other signaling protocols to register own MAC and IP addresses with call server. In the reply, each phone will get unique call ID which allows it to make or receive VoIP call. Attacker tries to capture registration packets and replaces MAC address from the packets with own MAC address. It helps the rogue node to register with victim IP address which causes call intending for victim node will be forwarded to attacker. SIP registration hijacking allows burglars to track, block and manipulate voice traffic. As end node registration is based on TCP connection, attack will be prevented by implementing SSL/TLS security policies [17].

2.2.9 Malformed Packets

The hacker creates malicious packets and forwards them to nodes inside VoIP networks with the help of networking protocols. The target node processes those packets, causes open unnecessary ports and processes which degrade performance of the nodes to handle VoIP traffic. New patches and software will be installed to maintain the node up-to-date and shutdown the security holes which are vulnerable to attack. New generation firewalls must be installed to provide protection against vulnerable packets by filtering packets based on inbound rules, outbound rules and connection security rules. Geneiatakis et al. have succeeded in developing a framework that provides defense against malformed packets for VoIP infrastructure[18]. The detection mechanism is based on signature detection which consists of two parts. First one, general signature detection (e.g., SIP METHOD, SIP URI, HEADERS) applicable to all the packets and second one is method specific (e.g., CALL-ID, Content-Type, INVITE _METHOD) differ from packets to packets.

2.2.10 SIP Message Modification

In message modification attack, by running network sniffing tools (e.g.,Wireshark), attacker penetrates traffic and tries to modify signaling message for better control over the VoIP network. Suppose a user initiates a call to victim's phone by sending SIP message to call server. Modification of SIP messages confuses and forces the server to connect rogue phone. User knows that he is connected to one user but actually the traffic is routed to attacker. SIP message modification is carried out by performing MITM attack such as MAC spoofing, IP spoofing or ARP poisoning. As SIP and RTP packets transmission are taken place over TCP and UDP connection; VoIP traffic must be encrypted by implementing SSL/TLS to prevent this attack[17].

2.2.11 SIP Cancel/ Bye Attack

Host (zombie) must be configured in promiscuous mode to lunch attack into VoIP network by sending SIP Cancel or Bye packets. Abnormal packets are created and sent to an IP phone from its connected IP phone by spoofing its IP address which will proceed to terminate the ongoing call. Attacker can perform this attack continuously for certain period of time by spoofing more than one IP addresses which causes denial of service attack. As both signaling and transport protocols use no authentication prior to data transmission, so, this attack can be prevented by encrypt the communication channels. Second, provide authentication between end device and call server and at last verification of authenticity of signaling message by end devices before processing [19].

2.2.12 SIP Malformed Command

In web based VoIP communication (e.g. Facebook, Google Hangout), Hyper Text Markup Language (HTML) plays a major role as it carries all the signaling information/ command in its body. Parsing SIP command within HTML code for all possible input is really a headache. Attacker tries to inject malformed SIP command in input field and send to server for processing as like SQL injection. In response either it breaks the server authentication or degrades the performance of server and end devices. In counter measure, whether packets are coming from genuine user or not will be confirmed by call server by verifying authenticity of SIP message before processing. Dictionary and fuzzy tests must be performed on HTML code that filtered tricky SIP malformed packets used to exploit server. M. Su and C. Tsai propose two functions to resists malformed SIP packets and flooding attack on call servers[20]. First function filters malformed packets and second one uses Chi-square test to measure flooding attack on SIP server.

2.2.13 SIP Redirect

Call server cache maintains data structure of Phone's caller ID, corresponding MAC and IP address. Attacker manipulates call server cache to confuse the call server for call redirection. So, SIP packets coming for receiver are redirected to attacker specified number. Attacker can perform DoS and DDoS attack by redirecting a single call or all the calls to void device(s). So, call server must be strong password protected and SIP must be authenticated to prevent redirection attack[19].

2.2.14 RTP Payload

Captured packets will be played later to listening the conversation between the end users using sniffing tools. Attacker can insert own voice inside RTP payload which degrade the quality of conversation and sometime changed in the meaning of conversation. In RTP tampering, header fields (sequence number, synchronization source Identifier, payload type, timestamp etc.) are tampered which make the packets either unusable or delayed, causes rejection at receiver end. In RTP redirection, header field of packets are modified with other receiver caller id and IP address causes packets intending for one will go to other. It can be prevented by configuring VoIP network with Secure Real-Time Transport Protocol (SRTP) instead of RTP[21]. It will encrypt the RTP packets propagate between callers.

2.2.15 Buffer Overflow

Buffer is the temporary storage allocated by OS in physical memory for processing data by computer program. Buffer is mainly divided into four types such as code, data, stack and heap segments. Attacker tries to perform buffer overflow attack by targeting at least one of the segments. It helps to steal or modify the sensitive information or install malicious code and execute it. Buffer overflow attacks are mainly executed by four ways such as long jump, function activation record, pointer subterfuge and malicious code execution. It can be defended by writing secure code, performing bound checking or static and dynamic code analysis and runtime code instrumentation[22].

2.2.16 Operating System

In VoIP communication network, IP phones, Call server, TFTP server, gateway and DHCP server etc. requires OS (e.g., Windows, Linux, Mac) to run. So, vulnerabilities in OS make them vulnerable[23]. OS vulnerabilities in VoIP phones are mainly of two types. Hard phones have in build embedded OS which is less vulnerable and more protected than soft phones. VoIP soft phones are software packages which are installed on computers connected to data network. Old hardware, unsupported drivers, bad integration of APIs, unsecure administrator APIs expose OS to attack. Like IP phones, web server OS, DHCP server, and call manager can be exploited by attacker for seeking of sensitive and crucial information (e.g., password, IP table, VoIP configuration file). As default configuration of OS is not secure, it is exposed to malwares to install. Its execution opens well known ports which helps attacker to run abnormal processes (e.g., free call, toll fraud). It can be pre-empted by hardening OS[24].

2.2.17 Malwares

A vulnerable piece of executable codes or program used by unknown third party to install in VoIP network and bring down its performance by hook or crook. Malicious programs or malwares are mainly classified as two categories, first one simple malwares and second one is self-replicated malwares[25]. Logic bomb and Trojan horse are come under non self-replicated/simple malware. Self-replicating malware such as virus and worm, who spread its infection over the network within few hours or days. Trojan horses are dispatched over network for remote control over victim VoIP phones. Logic bomb helps the attacker to trigger other dangerous attacks (DoS, DDoS, sniffing etc.) in timely manner. It will be prevented by installing updated antivirus and patching up VoIP system software on regular basis.

2.2.18 Application Flaws

As most of the VoIP communications are web based, it's vulnerable to two major application flaws such as Structured Query Language (SQL) Injection attack and cross site scripting attacks. In SQL Injection attack, malicious commands are inserted in SQL statements to gain unauthorized access to server database. It can be prevented by implementing three primary defense mechanisms such as defensive coding, SQL injection vulnerabilities detection and runtime SQL injection attack prevention[26]. In cross site scripting attack, hacker uses the advantages of scripting languages to launch attack by injecting malicious code inside the web application. It can be prevented by configuring strong authentication and validation for web based VoIP application[27].

2.2.19 TFTP Server Insertion

Hacker tries to plant rouge TFTP server in the network by disabling/ spoofing actual TFTP server. It forces IP phones to receive wrong configuration information (e.g., Call ID, SIP server IP address and phone number) which may provoke bill fraud attack. It will be prevented by encrypting and authenticating the channel between IP phones and TFTP server using TLS/ SSL. N. N. Mohamed et al. suggested compression and encryption technique to secure TFTP packets[28]. For compression, lossless algorithm (e.g., Huffman coding) and for encryption, symmetric encryption algorithm (e.g., AES, 3-DES) is used. Diffie-Hellman Key Exchange algorithm is used for distribution of symmetric key between client and server.

2.2.20 DHCP Server Starvation

Attacker generates random MAC addresses and creates DHCP request for each MAC address. By flooding DHCP server with DHCP requests, consumes DHCP IP pool and to overflow later. It is to be continued until reserved IP addresses DHCP timers will be expired. Dinu and Togan proposed digital certificate based DHCP server authentication to stop DHCP server starvation attack[29]. It uses asymmetric key cryptography and digital certificates for DHCP server authentication and verifying DHCP response from it to prevent starvation.

2.3 Defense Mechanisms to Prevent Attacks

Defense mechanisms provide basic counter measures to prevent potential VoIP attacks explained above are broadly classified into twelve types and listed in Table 2 [7,30].

2.3.1 Physical Access Control (PAC)

Physical securities can be implemented mainly three ways[31]. First, equipment should be placed and surrounded by multi-layer barriers, which will prevent from natural disasters like cyclone, floods etc. (e.g., wall, multiple locks, fireproof safes etc.). Second, deployment of surveillance systems such as smoke and heat detectors, cameras, alarms that decreases occurrences of manmade disasters with maximum amount. At last, practices must be implemented to prevent before any attack has been occur and fast recovery from damages, if any attack has occurred.

2.3.2 ARP Cache Protection (ACP)

Static ARP cache entries allow maintaining manual mapping between IP address to MAC address so that

Table 2. Defense mechanisms against attacks

Attacks / Defense Mechanisms	1.Physical Attacks	2. MAC Spoofing	3. ARP Spoofing	4. IP Spoofing	5. ICMP Flood	6. TCP/UDP Floods	7. TCP/ UDP Replay	8. SIP Registration Hijacking	9. Malformed Packets	10. SIP Message Modification	11. SIP Cancel/ Bye Attack	12. SIP Malformed Commands	13. SIP Redirect	14. RTP Payload	15. Buffer Overflow	16. Operating System	17. Malwares	18. Application Flaws	19. TFTP Server Insertion	20. DHCP server Starvation
PAC	√																			
ACP				√																
OSP														√		√	√	√		
PA		√	√	√	√	√	√	√	√	√	√	√	√	√	√			√	√	√
RC				√	√															
FC						√			√									√		
SVDT		√	√	√	√	√	√	√	√	√	√	√	√	√	√				√	√
CA																√			√	
SA							√	√		√	√	√	√		√					
ME							√													
IDS				√	√	√	√	√	√	√	√	√	√		√		√	√	√	√
Honeypot		√		√																

reply packets are discarded. ARP anti-spoofing software ignores ARP spoofing packets by it certification or crosschecking of its responses. It can be integrated with Dynamic Host Configuration Protocol (DHCP) server, so that all static and dynamic IP addresses are certified before used. Operating system security is provided by configuring its registry files that prevent burglar to update ARP cache. Registries files are found under HKEY_LOCAL_MACHINE\SYSTEM folder. Yang, Yang and Ding proposed a WinPcap driven system that monitors all ARP packets for ARP spoofing[32]. WinPcap driver captures and monitors the packets to verify whether the IP-MAC mapping is legal or not by checking packets coming from legal hosts, before update the cache.

2.3.3 Operating System (OS) Protection (OSP)

OS protection requires vulnerability assessment and management techniques such as patching OS holes, OS hardening, updating security software, OS auditing, proper ***priviligation*** to user accounts etc. Kaczmarek and Wrobel proposed integrity checking and recovery (ICAR) protection model comprises of three layers and it's responsible for hash creation, verification and configuration of security policies[33]. Data layer consists of sensitive information and database that contains hashes and backup of highly sensitive information. Kernel layer manages verification of data integrity, authenticity and confidentiality. Utility layer is responsible for configuring security policies and controlling the host.

2.3.4 Port Authentication (PA)

Strong port authentication can provide defense against interception, interruption and modification of traffic, interoperability between old and new network protocols and prevent malicious software execution. IEEE 802.1X port based authentication supplies security credentials such as user id, password or digital certificate to legal user[34]. User has to use the credential for verification, before access the resources. If server verifies the credentials are valid, user is permitted to utilize the resources. deGraaf, Aycock and Jacobson explained port knocking where authentication data is communicated throughout network ports to prevent unauthorized access[35].

2.3.5 Router Configuration (RC)

Fraudulent route update packets are blocked by configuring neighbor configuration. The configuration is available in the following routing protocols such as Boarder Gateway Protocol (BGP), DRP Server Agent, Intermediate System-Intermediate System (IS-IS) etc. Sehgal and Nath proposed secure routing protocol which has two phases[36]. In route discovery phase, Source node initiates a request to discover a route from source to destination. Route unitization phase, source selects one of the route and the destination has to confirm that route by sending reply packet to source. Source sends predecessor packet to notify intermediate nodes on the route that they should anticipate certain amount of data within a given time. When this packet reaches at destination, source receives an acknowledgement from destination. If not, there is a malicious user in between the path.

2.3.6 Firewall Configuration (FC)

Firewalls mainly configured into three types such as packet filters, stateful inspection and application proxy. Packet filter monitors all the packets header fields such as source and destination address, port number and protocol used based on predefined rules. Stateful inspection firewall tracks network connection state and differentiate packets based on the type of connection. It is also called dynamic packet filtering firewall and its fail to examine the content of the payload. Application proxy firewall allows the entire traffic pass through a proxy server, which verifies packets header including its content at application level for any malicious activities. Chacon, Benhaddou and Gurkan proposed Virtual Private Network (VPN) based firewall that provides more security to Boarder Gateway Router (BGR) by making voice information less vulnerable to both inside and outside attack[37].

2.3.7 Separate VoIP Data Traffic (SVDT)

In VoIP communication, both voice and data share common medium for transmission which raises threats against voice. Butcher, Li and Guo proposed separation of voice from normal data flow can block a number of attacks which are vulnerable to FTP, HTTP and SMTP etc. [7]. Separate physical network causes more expense, so, it can be possible through VLAN technology. VLANs are implemented by network switches allow routing on same VLAN between devices. Both VoIP voice and data are segmented using firewall where PCs are connected to data segment and VoIP phone are connected to voice segment.

2.3.8 Configuration Authentication (CA)

In VoIP telephony network obtaining the configuration information from unwanted vulnerable server makes end phones more sensitive to launch attacks. Danforth and Gould described different ways of authenticating TFTP server before downloading configuration file[38]. During manufacturing time, the VoIP phones are configured with public keys of different TFTP servers gives a way for authentication. Another way the handsets are configured with a key (public or secret) of TFTP server. After getting own IP address and TFTP server IP address from DHCP server, phone should establish a secure connection with TFTP server using SSL/TLS. During handshaking, verification happens using public key that phone contains and private key of TFTP server.

2.3.9 Signaling Authentication (SA)

In VoIP network, SIP is used to establish, redirect or terminate the connection. Internet Protocol Security (IPSec) and SSL/TLS are used to provide strong authentication and encryption against attack. Key agreement protocol is used in small scale deployment where trust being established between phones and server. Kilinc and Yanik presented different authentication and key management schemes for SIP protocols which mainly include Password Authenticated Key Exchange (PAKE) based schemes, Hash and Symmetric Encryption based schemes, Public Key Cryptography (PKC) schemes and ID Based and Weil Pairing based schemes[39].

2.3.10 Medium Encryption (ME)

In VoIP, medium encryption is broadly classified into two types such as symmetric and asymmetric encryption[40]. Encryption strength is mainly depending upon the strength of the algorithm and the size of key is used. In encryption techniques key management and distribution also play an important role[41]. To maintain confidentiality and integrity, symmetric key is distributed with the help of the techniques such as simple secret key distribution, secret key distribution with confidentiality and authentication and hybrid key distribution. Public key sharing has taken place using public announcement, publicly available directory, public key authority or public key certificate techniques.

2.3.11 Intrusion Detection System (IDS)

Hardware/ software are used to monitor network traffic for malicious and unlawful actions and notify to admin by warning message or raising alarm. Basically IDS is of two types such as Network Intrusion Detection System and Host Intrusion Detection System. It commonly uses three methodologies to track down malicious activities. Signature based detection is the valuable and straight forward methods to uncover known threats. Anomaly based detection is effective one to identify new and expected threats. Stateful protocol analysis tracks down protocols performance and differentiates abnormal flow of commands. All these three methods use five different approaches such as statistics based, pattern based, rule based, state based and heuristic based which were discussed by Liao, Lin, Lin and Tung[42].

2.3.12 Honeypot

It is a trap set to detect, deflect or counter attempt at unauthorized use of information system which is seems to be contain of information or resources of values to attackers. It is classified based on their deployment and based on their level of involvement[43]. Based on its deployment it's of two types, one is production honeypot and second one is research honeypot. According to level of involvement or design perspective, pure honeypot, higher interaction honeypots, low interaction honeypots are three types of honeypots. Goel, Sardana and Joshi presented a wide range of honeypot systems and proposed framework for honeypot system that enclose a broad range of honeypot architectures and categories previous systems according to framework based on attacks[44].

3. Proposed Work

VoIP cloud with all security configurations as discuss earlier is not enough to provide security against all the threats discussed earlier. So, effective network based IPS architecture using LVD technique is proposed for VoIP cloud shown in Figure 3.

The proposed system involves twosteps process. 1) Video watermarking and extraction and 2) Verification using LVD system

3.1 Video Watermarking and Extraction

There are several VoIP-specific protocols but they fall in two categories: (i) transport protocols (e.g., RTP) and (ii) signaling protocols (e.g., SIP). Transport protocols carry the live video data after the proper the connection. Signaling protocols executes control information like CONNECT, DISCONNECT etc.

Sequence numbers play important roles to prevent SSL attacks in real time communication. So at the first step the sequence number is attached with the video/ frames taken during the connection/disconnection time. Since the control information is not a video data, it has to be embedded with video using watermark embedding technique and forwarded to the server side for authentication as illustrated in Figure 4. In the server side, the watermarked video will be extracted and the video submitted to the LVD for verification.

IPS in VoIP cloud not only monitors the packets are coming from legitimate caller or not but also monitors sequence number of packets and its content. In real time communication succeeding packet sequence number always larger than current packet sequence number and contents varies packet to packet and time to time which helps IPS to prevent SSL attacks to be taken place. If any packet with old sequence number or repetition of information inside the payload, it will be dropped by VoIP server. Always signaling packets will get high priority over data packets.

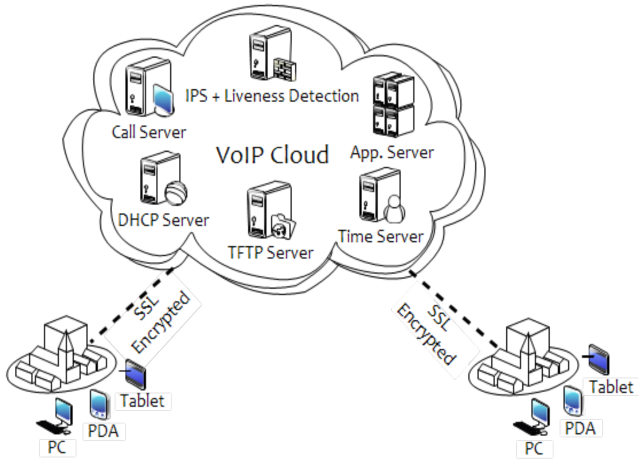

Figure 3. Proposed VoIP cloud architecture.

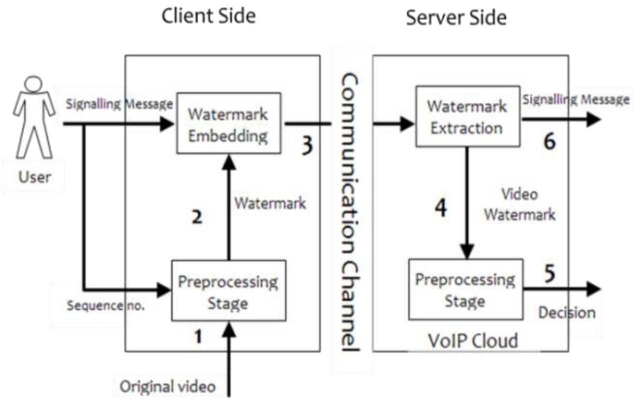

Figure 4. Video watermarking scheme for signaling message.

3.2 Verification using LVD System

The aim of the LVD is to determine if the biometric data is being captured whether it's from a legitimate live user or it is replayed or synthetic. It is verified by correlating the user's voice with lip-face motion. The Whole LVD is divided into seven steps. The whole process of verification is divided into two stages. The flow diagram of LVD is shown in Figure 5. There are two types of videos passed to LVD system for verification. 1) Extracted watermarked video2) RTP message along with the video.

3.2.1 *Seven Steps of LVD*

Step 1: Image and Voice Acquisition: Video and Voice are captured by web cam and microphone respectively.

Step 2: Image and Speech Enhancement: The captured information goes for reduction of noise, and this can be achieved by smoothing and sharpening the audio- visual data.

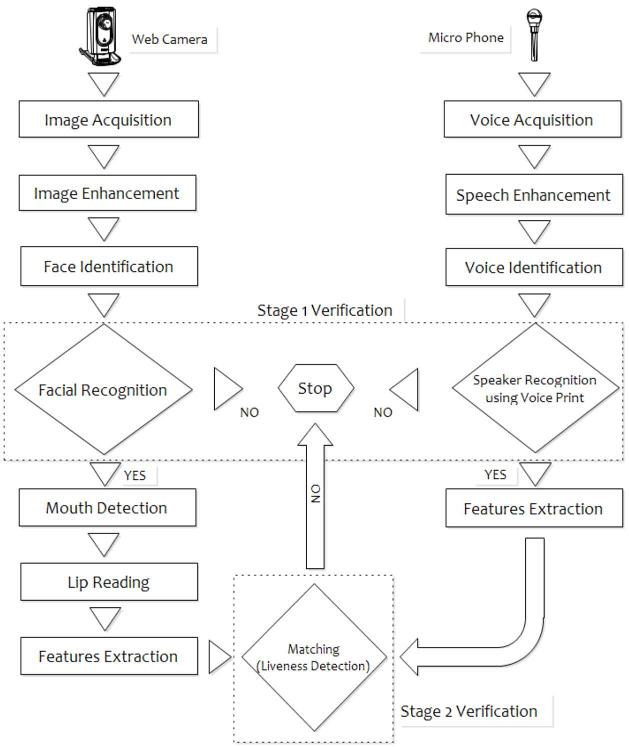

Figure 5. Flow diagram of liveness voice detection technique.

Step 3: Face and Voice Identification: Real time faces detection and speech detection can be done by motion analysis. Two level audio- visual fusion techniques can be used for effective face and voice identification by localization of multiple active speakers[45]. First level is based on speaker activity detection used to find out who are the live speakers and second level uses Gaussian method for integration of audio-visual modalities results to increase robustness.

Step 4: Facial Verification and Speaker Recognition: In Face verification and speaker recognition process, user authenticity is checked. If NO, the call is terminated. If yes, it will go for next level verification.

Step 5: Features Extraction: From lip movement, the motion based features will be extracted *(Phase 1)*. From speech, the corresponding speech features will be extracted *(Phase 2)*. Extracted features at both phases must be language and text independent. Lip movement involves the following two elements. *Fastness:* frequency is the prime factor calculated from lip motion. *Loudness:* Power is the prime factor calculated from area covered by lips (distance between top and bottom lip).

Step 6: Liveness Voice Detection: Features from phase 1 is correlated with features from phase 2. If it falls within fixed threshold values range, it is Ok, the connection will be continued else it will be terminated. In LVD, liveness score evolution algorithm can be used to measure the synchrony between the lip movement and voice in video sequence[46]. Multimodal system based on cross modal fusion technique can be also used for liveness detection[47]. Audio and visual speech features are extracted from video sequence to measure the degree of synchrony between the lip movement and voice in video sequence.

3.3 IPS Deployment

As VoIP communication handles real time data, IPS deployment is also an important factor to make the service more effective and efficient. In pass-by monitoring configuration, a copy of the traffic is sent to the IPS while the original packet travels to one Public Network to another public network as shown in Figure 6. If the IPS identifies an anomaly with the packet, the IDS/IPS can either log/record the activity or prevent the attack from being successful. As an effective IPS, false positive/ negative should have minimized with maximum efficiency.

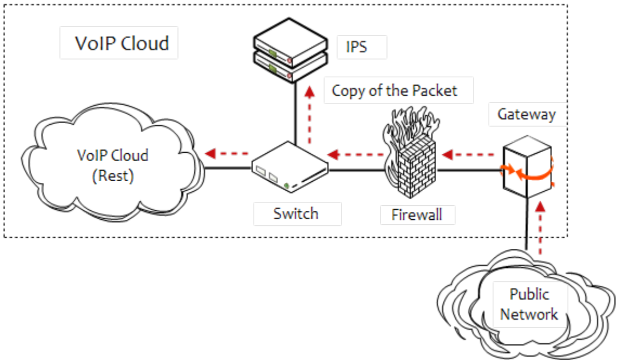

Figure 6. Proposed N-IPS architecture.

4. Conclusion

The flexible and reliable communication is reached by sending the voice over the internet by using new generation VoIP services. This research paper summarizes security threats related to VoIP cloud. Due to VoIP vulnerabilities including eavesdropping, DoS, D-DoS, MITM attack, it is necessary to protect signaling and real time information. A security framework for VoIP cloud is proposed which uses the concept of LVD to provide resistance against threats. The use of pass-by IPS makes this proposed framework has no effect on quality of VoIP calls as the copy of the original packets are forwarded to IPS and analyzed. At last, further research has to be performed to raise the level of security due to randomness of occurrence of cyber-attacks.

5. References

1. Hartpence B. Introduction to voice over the internet protocol. Packet Guide to Voice over IP. Oram A, Gulick M, editors. O'Reilly: Sebastopol, CA; 2013.
2. Devi GU, Kaushik KV, Sreeveer B, Prasad KS. VoIP over Mobile Wi-Fi hotspot. Indian Journal of Science and Technology. 2015 Jan; 8(S2):195–9. DOI: 10.17485/ijst/2015/v8iS2/58751.
3. Patinge SA, Soni PD. Metamorphosis in VoIP cloud computing services used in VoIP. International Journal of Application Innovation in Engineering Management. 2013; 2(2):236–9.
4. Mahmood Z. Cloud computing: characteristics and deployment approaches. 11th IEEE International Conference Computer and Information Technology (CIT); Pafos: Cyprus; 2011. p. 121–6.
5. Shyamala K, Rani TS. An analysis on efficient resource allocation mechanisms in cloud computing. Indian Journal

of Science and Technology. 2015 May; 8(9):814–21. DOI: 10.17485/ijst/2015/v8i9/50180.

6. Kim D. A survey of balloon networking applications and technologies. Available from: http://www.cse.wustl.edu/~jain/cse570-13/ftp/balloon/index.html. [Cited 2014 Aug].
7. Butcher D, Li X, Guo J. Security challenge and defense in VoIP infrastructures. IEEE Transactions on Systems Man and Cybernetics Part C: Applications Reviews. 2007; 37(6):1152–62.
8. Graves K. Certified ethical hacker study guide, 4th ed. Wiley: Danvers, MA; 2010.
9. Sawada H, Higashimoto T. A mechanical voice system and its adaptive learning for the mimicry of human vocalization. Proceedings IEEE International Symposium on Computational Intelligence Robotics and Automation; Cobe, Japan; 2003. p. 1040–45.
10. Hatkar AA, Varade GA, Hatkar AP. Media access control spoofing techniques and counter measures. International Journal Scientific & Engineering Research. 2012; 2(6):1–5 .
11. Nam SY, Kim D, Kim J. Enhanced ARP: preventing ARP poisoning-based Man-in-the-Middle Attacks. IEEE Communications Letters. 2010; 14(2):187–9.
12. Ma Y. An effective method for defense against IP spoofing attack. IEEE 6th International Conference on Wireless Communications Networking and Mobile Computing (WiCOM); Chengdu: China; 2010. p. 1–4.
13. Barbhuiya FA, Roopa S, Ratti R, Biswas S, Nandi S. An active detection mechanism for detecting ICMP based attacks. IEEE 11th International Conference on Trust, Security and Privacy in Computing and Communications; Liverpool: England; 2012. p. 51–58.
14. Haris SHC, Ahmad RB, Ghani MAHA, Wal GM. TCP SYN flood detection based on payload analysis. Proceedings IEEE Student Conference on Research and Development (SCOReD); Putrajaya: Malasia; 2010. p. 149–53.
15. Bardas AG, Zomlot L, Sundaramurthy SC. Classification of UDP traffic for DDoS detection. USENIX 5th International Workshop on Large-Scale Exploits and Emergent Threats (LEET); San Jose: CA; 2012. p. 1–8.
16. Ali FHM, Yunos R, Alias MAM. Simple port knocking method against TCP replay attack and port scanning. IEEE International Conference on Cyber Security. Cyber Warfare and Digital Forensic (CyberSec); Kuala Lumpur: Malasia; 2012. p. 247–52.
17. Stalling W. Transport-level security. Cryptography and Network Security. Horton M, editor, 5th ed., Pearson: Upper Saddle River, NJ; 2011. p. 485–20.
18. Geneiatakis D, Kambourakis G, Lambrinoudakis C, Dagiuklas T, Gritzalis S. A frame for protecting a SIP-based infrastructure against malformed message attacks. Computer Network. 2007; 51(10):2580–93.
19. Zhang G, Pallares JJ, Rebahi Y, Fischer-Hubner S. SIP proxies: New reflectors in the internet. Communications Multimedia Security; Springer : Verlag Heidelberg; 2010.
20. SuM Y, Tsai CH. An approach to resisting malformed and flooding attacks on SIP servers. Journal of Networks. 2015; 10(2):77–84.
21. Hartpence B. The real-time transport protocol and the real-time control protocol. Packet Guide to Voice over IP, Oram A, Gulik M, editors, 1st ed.; O'Reilly: Sebastopol, CA; 2013.
22. Fu D, Shi F. Buffer overflow exploit and defensive techniques. IEEE International Conference on Multimedia Information Networking and Security (MINES); Nanjing, China; 2012. p. 87–90.
23. Ransome JF, Rittinghouse JR. VoIP security risks. VoIP Security, Casey E, editor; Elsevier: Burlington, MA; 2005.
24. Hardening the operating system. Available from: http://cdn.ttgtmedia.com/searchEnterpriseLinux/downloads/466_HTC_Linux_02.pdf. [Citied 2014 Oct].
25. Filiol E. Taxonomy, techniques and tools. Computer Viruses: From Theory to Applications, 1st ed.; Springer: Verlag, France; 2004.
26. Shar LK, Tan HBK. Defeating SQL injection. IEEE Computer: Gender Diversity in Computing. 2013; 46(3):69–77.
27. Natan RB. Application security. Implementing Database Security and Auditing; Elsevier: Burlington, MA; 2005.
28. Mohamed NN, Mashim H, Yussoff YM. Compression and encryption technique on securing TFTP packet. IEEE Symposium on Computer Application Industrial Electronics (ISCAIE); Penang, Malaysia; 2014. p. 198–202.
29. Dinu DD, Togan M. DHCP server authentication using digital certificates. Proceedings IEEE 10th International Conference Communications (COMM); Bucharest, Romania; 2014. p. 1–6.
30. Keromytis AD. A comprehensive survey of Voice over IP security research. IEEE Communications Surveys & Tutorials. 2012; 14(2):514–37.
31. Graves K. Physical site security. Certified Ethical Hacker Study Guide, Parsons K, Carson C, 4th ed.; Wiley: Danvers, MA; 2010.
32. Yang M, WangY, Ding H. Design of WinPcap based ARP spoofing defense system. IEEE 4th International Conference on Instrumentation Measurement Computer, Communication Control (IMCCC); Harbin, Heilongjiang; 2014. p. 221–5.
33. Kaczmarek J, Wrobel MR. Operating system security by integrity checking and recovery using write-protected storage. IET Information Security. 2014; 8(2):122–31.

34. IEEE 802.1X port-based authentication, CISCO. Available from: http://www.cisco.com/c/en/us/td/docs/switches/lan/catalyst6500/ios/122SX/configuration/guide/book/dot1x.html#wp1133592. [Cited 2015 Jan].
35. deGraaf R, Aycock J, Jacobson M. Improved port knocking with strong authentication. Proceeding. IEEE 21st Annual Computer Security Applications Conference (ACSAC); Tucson, Arizona; 2005. p. 453–62.
36. Sehgal PK, Nath R. An encryption based dynamic and secure routing protocol for mobile Ad-hoc network. International Journal of Computer Science Security. 2009; 3(1):16–22.
37. Chacon S, Benhaddou D, Gurkan D. Secure Voice over Internet Protocol (VoIP) using Virtual Private Networks (VPN) and Internet Protocol Security (IPSec). IEEE Region 5 Tech. Professional and Student Conference (TPSC); San Antonio: TX; 2006. p. 218–22.
38. Danforth A, Gould K. Method to block unauthorized access to TFTP server configuration files, U.S. Patent 7293282 B2, 2007.
39. Kilinc HH, Yanik T. A survey of SIP authentication and key agreement schemes. IEEE Communications Survey and Tutorials. 2014;16(2):1005–23.
40. Anderson R. Cryptography. Security Engineering: A Guide to Building Dependable Distributed Systems, Long C, 2nd ed., Wiley: Indianapolis, IN; 2008. p. 73–14.
41. Stalling W. Key management and distribution. Cryptography and Network Security. Horton M, 5th ed., Pearson: Upper Saddle River, NJ; 2011. p. 410–43.
42. Liao HJ, Lin CHR, Lin YC, Tung KY. Intrusion detection system: A comprehensive review. Journal of Network and Computer Applications. 2013; 36(1):16–24.
43. Audiopedia. Honeypot (computing). Available from: https://www.youtube.com/watch?v=2fXAw33jOBk. [Cited 2014 Dec].
44. Goel R, Sardana A, Joshi RC. Wireless honeypot: framework, architectures and tools. International Journal of Network Security. 2013; 15(5):373–83.
45. Li Z, Grochulla M, Thormahlen T. Multiple active speaker localization based on audio-visual fusion in two stages. Proceedings IEEE International Conference on Multisensor Fusion Integration Intelligence Systems (MFI); Hamburg: Germany; 2012. p. 262–68.
46. Zhu ZY, He QH, Feng XH, Xiongli Y, Wang ZF. Liveness detection using time drift between lip movement and voice. Proceedings IEEE International Conference on Machine Learning Cybernetics (ICMLC); Tianjin: China; 2013. p. 973–78.
47. Chetty G. Biometric liveness detection based on cross modal fusion. IEEE 12th International Conference on Information Fusion (FUSION). Seattle: WA; 2009. p. 2255–62.