Security in Voice over IP Networks

Juan C Pelaez, M.C.S.
Florida Atlantic University, 777 Glades Road SE-300, Boca Raton, FL 33431-0991 USA
jpelaez@ieee.org,

Eduardo B. Fernandez, PhD
Florida Atlantic University, 777 Glades Road SE-408, Boca Raton, FL 33431-0991 USA
ed@cse.fau.edu

Abstract
Voice over IP (VoIP) has had a strong effect on global communications by allowing human voice and fax information to travel over existing packet data networks along with traditional data packets. The convergence of voice and data in the same network brings both benefits and constraints to users. Among the several issues that need to be addressed when deploying this technology, security is one of the most critical. We give an overview of VoIP and its applications including its internetworking with wireless networks and provide UML models of some aspects of its infrastructure. Finally, we present some object-oriented security patterns based on a systematic analysis of attacks against a VoIP network and the existing techniques to mitigate these attacks.

Keywords
IP Protocol, networks, object-oriented patterns, security, VoIP.

1. Introduction

VoIP is defined as the transport of voice over IP based networks. Any data network that uses IP can be used to establish this service. VoIP uses IP to transmit voice as packets over an IP network. Therefore, VoIP can be achieved on any data network that uses IP, such as the Internet, intranets and Local Area Networks (LAN), where digitized voice packets are transmitted over the IP network.

IP Telephony enables the transfer of voice data over a packet-switched network as opposed to the traditional circuit-switched networks of today's telephone companies. VoIP can be considered as one more transport technique within the IP layer.

Existing network infrastructures can be used to carry both data and voice traffic, which is very attractive to new users. Savings come from eliminating the need to purchase new Private Branch Exchange (PBX) equipment, and from reducing staff and maintenance costs, as only one network needs to be supported

[Wei01]. The possible savings from the cost of long distance per minute charges of sending voice traffic via existing carriers provide extra incentives for moving to VoIP.

The transmission of VoIP networks enables a wide variety of applications, and VoIP can be applied to almost any voice communications requirement, ranging from a simple inter-office intercom system to complex multi-point teleconferencing/shared screen environments.

In VoIP, in addition to delivering voice, the IP protocol performs some of the related functions of the voice network which are necessary to convert the whole network into a full system. Some of these functions include special features, collect calling, gateways into the public voice network, and associated actions.

The purpose of this paper is to systematically define the attacks against a VoIP network in order to select the techniques that could mitigate these attacks. This paper will address some of the most important existing VoIP security issues, and will give a detailed presentation of problems which exist or are likely to exist in the future. To effectively analyze critical security issues in VoIP networks, we start with an overview of VoIP in Section 2; Section 3 models the actual IP telephony infrastructure. To develop object oriented patterns, we analyze the attacks against the VoIP infrastructure from the H.323 and Session Initiation Protocol (SIP) standards perspective in Section 4. Section 5 presents some security architectures while Section 6 provides some conclusions.

2. VoIP Overview

There are three different types of connections for setting up a call. In all the cases, the IP Protocol is used: (1) PC-to-PC, where individuals talk online through their PCs, (2) PC-to-Telephone, where individuals make and receive voice calls and messages while on the Internet, and (3) Telephone-to-Telephone, where calls are made and receive using regular phones connected to Public Switched Telephone Network (PSTN) or IP-telephones connected to a data net.

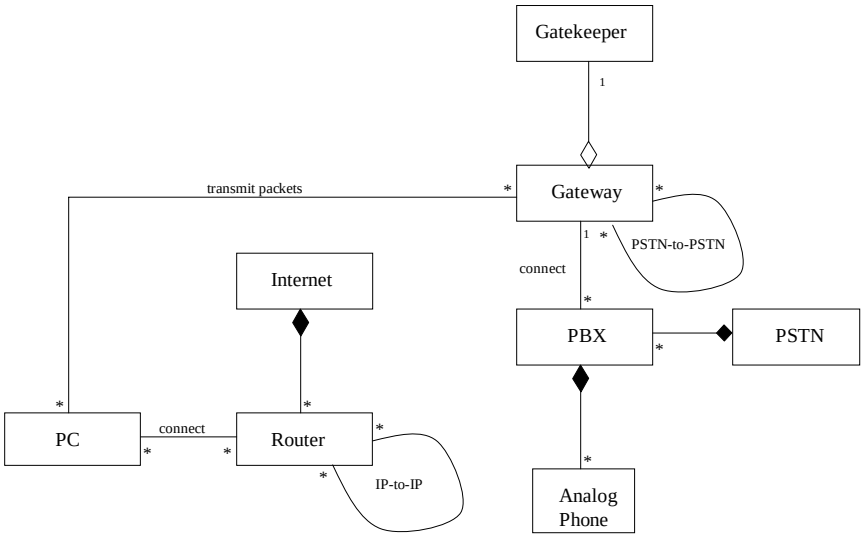

Figure 1 Class Diagram for a simplified VoIP architecture

Figure 1 shows a class diagram describing how an IP telephony system integrates with the PSTN. This model shows how it becomes possible to place a call from a regular telephone number to a PC running an

H.323 client (and vice versa). The PBX which supports the standard phone (caller), formats caller and callee numbers and forwards them to the VoIP gateway via PSTN network. The gateway takes the voice call from circuit-switched PSTN and places it on the IP network. The gateway queries the gatekeeper via the IP network with caller/callee numbers (note that the Voice packets do not go through the gatekeeper; only the call signaling goes through it) and the gatekeeper translates into a routing number based upon service logic. Finally the gateway routes the call to called party (callee).

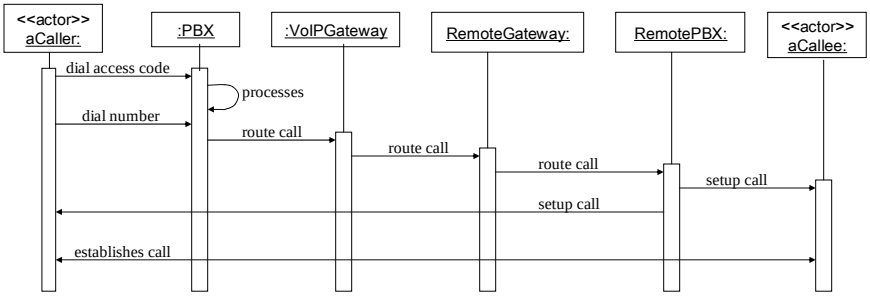

Figure 2 Sequence diagram for a telephone-to-telephone connection

Figure 2 shows a sequence diagram for a telephone-to-telephone connection type. In this case, the model presented in Figure 1 could be altered to have analog telephones as endpoints, rather than PCs in order to provide all types of connections for setting up a call.

3. VoIP Protocols

VoIP uses the Real-Time Protocol (RTP) for transport, the Real-Time Transport Protocol (RTCP) for Quality of Service (QoS) and H.323, SIP, MGCP (Media Gateway Control Protocol/Megaco) for signaling. These protocols operate in the application layer; that is, on top of the IP protocol.

Most current VoIP implementations use the H.323 protocol, the same protocol used for IP video. Users prefer H.323 more than SIP, but this may be primarily due to the earlier release of H.323 (in the 90's) [Wei01].

3.1. The H.323 Family of Protocols

H.323 is a multimedia standard which provides a foundation for transporting voice, video and data communications in an IP-based network. The H.323 standard is a part of the H.32x protocol family, which includes, besides H.323, standards like H.324 (standard for multimedia transport over switched circuit networks) and H.320 (standard for ISDNs) among others [DSQ01]. This standard is specified by the ITU research group and was first approved in 1996. H.323 runs on top of TCP in layer 4, and uses TCP for call setup. Traffic is actually carried on the Real Time Protocol (RTP) which runs on top of the User Datagram Protocol (UDP) [Tan03].

H.323 defines four logical components, including terminals, gateways, gatekeepers and multipoint control units (MCU).

Terminals are used for real-time bidirectional multimedia communications. IP phones connected to LANs (a.k.a. hardphones) are included, as well as PC-based IP Phones (a.k.a. Softphones). Softphones are applications installed on user systems (e.g. desktops) with speakers and microphones which reside in the data segment (implemented by VLANs). On the other hand, Hardphones are located in VLANs that support only IP telephony services. All H.323 terminals have to support H.245 (control channel), Q.931 (required for call signaling and setting up the call), Registration Admission Status (RAS, used for interacting with the gatekeeper) and the Real Time Transport Protocol (RTP) [Jai00]. In addition to supporting audio communications, terminals can be used to support video or data communications.

As previously mentioned, the gateway is the interface between the PSTN and the Internet. A gateway provides translation of protocols for call setup and release, conversion of media formats between different networks, and the transfer of information between H.323 and non–H.323 networks.

Gatekeepers provide call-control services for H.323 endpoints, such as address translation, admission control, bandwidth management, zone-management, and call-routing services. They provide authentication services to allow end-users to register on the VoIP network. Although they are an optional entity in the H.323 environment [Mil02], Gatekeepers are in practice the focal point for all calls within the H.323 network.

An *MCU* provides support for multi-conferencing between three or more H.323 terminals [Wei01]. All terminals participating in the conference establish a connection with the MCU. The MCU manages conference resources, negotiates between terminals for the purpose of determining the audio or video coder/decoder to use, and may handle the media stream.

In a VoIP environment, PBXs are replaced by server-based IP PBXs. These servers act like call processing managers providing call setup and routing the calls throughout the network to other voice devices. A class diagram for VoIP components is shown in Figure 3. The layer 2 QoS enabled switch provides connectivity and network availability between H.323 components.

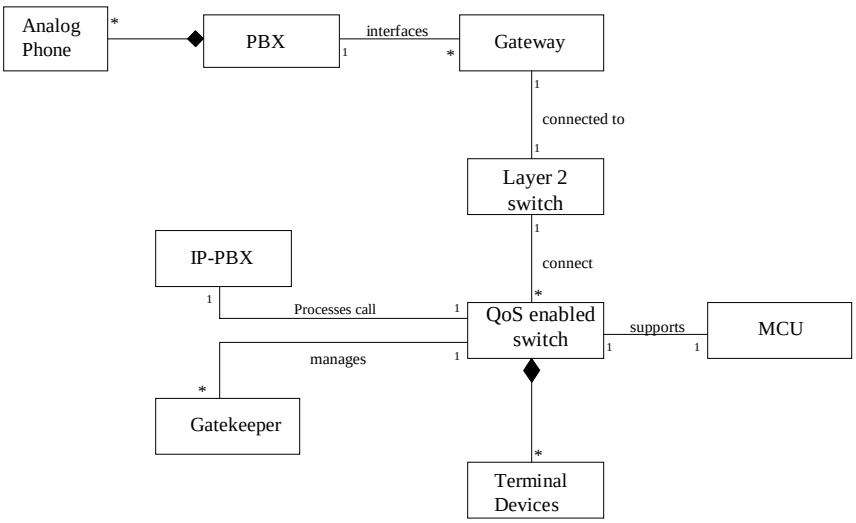

Figure 3 Class Diagram for a H.323 architecture

3.2. The Session Initiation Protocol

Session Initiation Protocol (SIP) is the IETF's standard for multimedia conferencing over IP. SIP is an application-layer control (signaling) protocol used for creating, modifying and terminating sessions with one or more participants. These sessions can include Internet multimedia conferences, Internet telephone calls and multimedia distribution. SIP is a less complicated protocol, and it is more flexible than H.323.

SIP is a text-based protocol, similar to HTTP and SMTP, for initiating interactive communication sessions between users. Such sessions may include voice, video, chat, interactive games, and virtual reality.

Signaling allows call information to be carried across network boundaries. Session management provides the ability to control the attributes of an end-to-end call.

SIP supports IP mobility for VoIP WLAN applications by providing handoff capabilities at the application layer. SIP can make direct use of Dynamic Host Control Protocol (DHCP) when connecting to an 802.11 AP for binding an IP address [Bas04].

3.2.1. SIP Architecture

The main components of SIP-based systems are user agents and servers:

User Agents (UAs), are combinations of User Agent Clients (UAC) and User Agent Servers (UAS). A UAC is responsible for initiating a call by sending a URL-addressed INVITE to the intended recipient. A UAS receives requests and sends back responses.

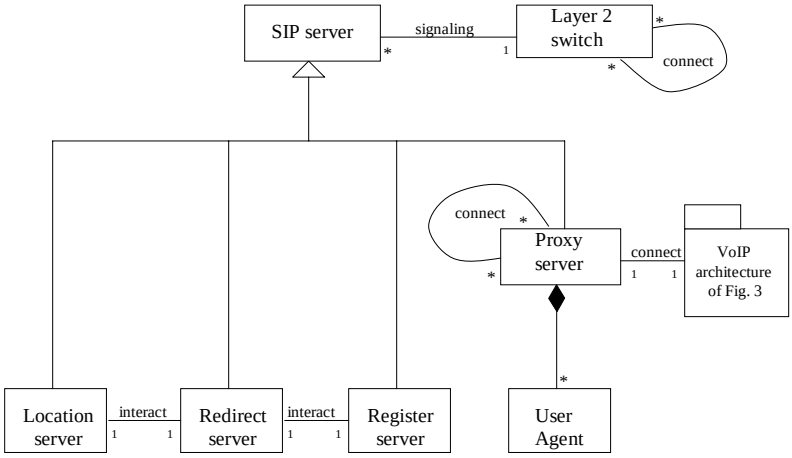

Figure 4 Class diagram of the SIP Architecture

Servers can be classified as:

Proxy servers operate on behalf of users and are responsible for routing and delivering messages.

Redirect servers, keep a user database, which allows them to inform proxy servers of a user location.

Location servers are used by a Redirect server or a Proxy server to obtain information about a called party's possible location.

Registrar servers, save information about where a party can be found.

Figure 4 shows the components for a SIP-based network. The proxy server is connected to a VoIP gateway (to make possible a call from a regular telephone to an IP phone) and to other proxy servers. The rest of the VoIP architecture is similar to Figure 3 and represented by a UML package. Once the call has been established, the RTP media streams flow between the end stations directly.

The SIP proxy server concept allows SIP to handle both firewall functions and network address translations (NAT), which are pervasive in home network topologies (assuming WLAN deployments on IPv4 networks). SIP will greatly benefit from the widespread use of IPv6, which would allow direct addressing of a mobile node client.

4. Possible Attacks to VoIP

4.1. Roles and Rights in a basic VoIP Model

In this section we will use Role Based Access Control and use cases to study the roles and rights of the human components in a VoIP network.

Because VoIP networks are vulnerable to attacks from external domains and internal sources, the human components of this system can be classified as described in section 4.1.1. We apply here the methodology of [Fer04], where attacks are related to use cases and role rights are derived from use cases.

4.1.1 VoIP Internal Roles

- **Internal subscriber** is a VoIP user such an employee. Internal subscribers are allowed to make and receive voice calls by either using standard or IP phones (hardphones and softphones). They also have access to data services by using terminal devices (e.g. PCs).
- **Network Administrator**. This role is responsible for maintaining a VoIP security network perimeter and routinely auditing the VoIP system in order to monitor user activities. The network security administrator is also responsible of properly configuring security mechanisms and reacting in the presence of attacks.
- **Network Auditor.** This role is responsible for performing audit logs to verify the integrity of the VoIP system. Auditing is especially useful for identifying potential security breaches or break-in attempts.
- **Network Operator** is responsible of protecting the system from being compromised, so that each voice call can be accounted to the appropriate user. He is also responsible for booting and shutting down the system, performing routine maintenance of servers, performing system performance metering and on-line tests, and in general responding to various relevant user requests.

4.1.2 VoIP External Roles

- **Remote subscriber** are VoIP system users such as employees who occasionally work from home. They are given access to voice and services only from their homes.
- **Law Enforcement Officer** refers to a legal agent who redirects duplicated media packets to law enforcement, for the purpose of wiretapping. The agent has access to corporate servers in order to intercept data and voice packets.

In addition to the mentioned roles, the simplified network model shown in Figure 1 will be used to systematically analyze the different types of attacks against the VoIP network.

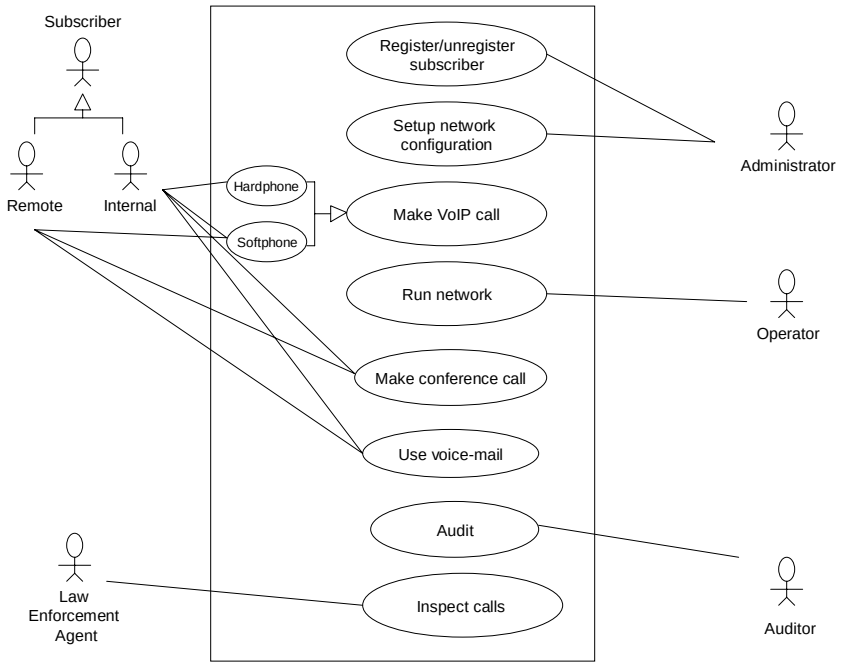

Figure 5. VoIP Use case Diagram

Figure 5 shows a Use Case diagram with the internal and external roles in a VoIP environment. From this diagram we can deduce the needed rights for each role. These roles and rights in a VoIP model are also shown in the Class diagram with authorizations in Figure 6.

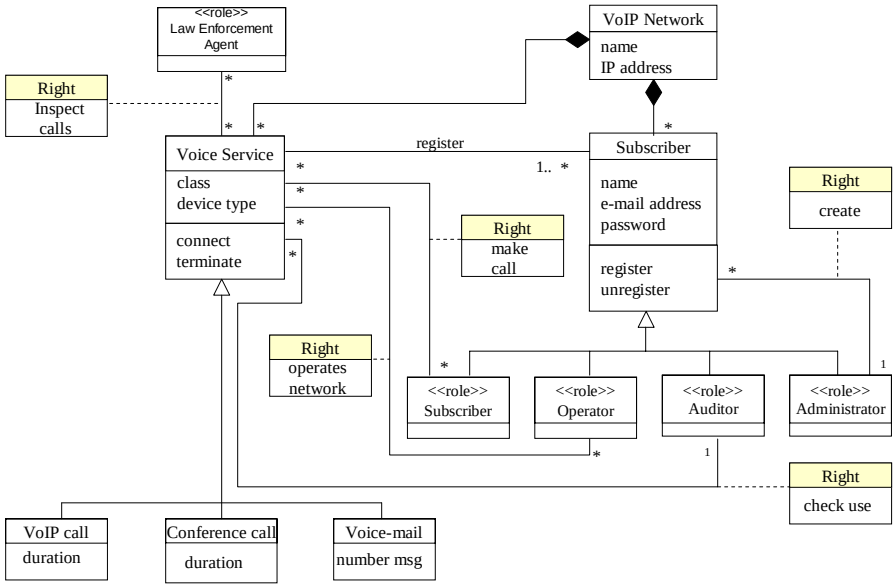

Figure 6. Class diagram with authorizations in VoIP

4.2. Attacks against the VoIP Network

Based on the Use Case Diagram of Figure 5, the attacks against the VoIP infrastructure will be discussed in the following sections.

4.2.1 Attacks when making/receiving a VoIP Call

Many of the already well-known security vulnerabilities in data networks can have an adverse impact on voice communications and need to be protected against [Pog03]; therefore VoIP users expect security. The attacks when making/receiving a voice call can be classified as follows:

Theft of service is the ability of a malicious user or intruder to place fraudulent calls. In this case the attacker simply wants to use a service without paying for it, so this attack is intended for the service provider.

There are numerous methods the hacker can use to accomplish this task. In a basic case of toll fraud, an unauthorized user places calls using an unattended IP phone or assuming the identity of the legitimate user of the telephone. In a more complex attack, a rogue IP phone may be placed on the network or a breached gateway may be used to make unauthorized calls. As mentioned before, gateways are used for routing packetized voice between the source and the destination within the IP network.

Masquerading, occurs when a hacker is able to trick a remote user into believing he is talking to his intended recipient when in fact he is really talking to the hacker. Such an attack typically occurs with the hacker assuming the identity of someone who is not well-known to the target. A masquerade attack usually includes one of the other forms of active attacks [Sta02]. A complex attack can be achieved by placing a rogue IP phone in the network and then assuming the identity of a valid IP phone via a secondary exploit.

IP Spoofing, occurs when a hacker inside or outside a network impersonates a trusted computer. There are two methods of doing this. The hacker can use either an IP address that is within the range of trusted IP addresses for a network or an authorized external trusted IP address that has access to specified resources on a network. As in Caller ID Spoofing attacks (i.e. Masquerading), IP spoofing attacks can be used to launch other types of attacks. A typical example is to launch a denial-of-service (DoS) attack using spoofed source addresses to hide the hacker's identity. Without some defense a hacker might be able to spoof the address of the IP-PBX and flood the entire voice segment with UDP packets.

With user identification based on the IP layer and the IP layer easily tampered with, it is easy for unauthorized users to impersonate legitimate ones by marking packets sent over these networks with a “borrowed” IP address. These abuses of services and benefits (e.g. making international calls) occur at the expense of legitimate users, who are often completely unsuspecting until the bill arrives—long after the abuser has disappeared [IEC01].

Call Interception is the unauthorized monitoring of voice packets or RTCP transmissions. Hackers could capture the packets and decode their voice packet payload as they traverse a large network. This kind of attack is the equivalent of wiretapping in a circuit-switched telephone system.

Due to the fact that voice travels in packets over the data network, hackers can use data-sniffing and other hacking tools to identify, modify, store and play back unprotected voice communications traversing the network, thus violating confidentiality.

A packet sniffer is a software application that uses a network adapter card in promiscuous mode (a mode in which the network adapter card sends all packets received on the physical network wire to an application for processing) to capture all network packets that are sent across a particular collision domain. This packet sniffer application can reside in a general-purpose computer attached, for example, in a corporate’s local area network [Fer05a]. For example, the tool "voice over misconfigured Internet

telephones" (a.k.a. “vomit”), takes an IP phone conversation trace captured by the UNIX tool tcpdump, and reassembles it into a wave file which makes listening easy [Pog03]. Figure 7 shows the sequence of steps that hackers use to monitor a VoIP conversation. With tcpdump, hackers can identify the IP and MAC address of the phone to be attacked By using an Address Resolution Protocol (ARP) spoofing tool, the attacker could impersonate the local gateway and the IP phone on the network, creating a default gateway [Pog03]. This allows IP traffic to and from the target IP phone to be monitored by the attacking workstation.

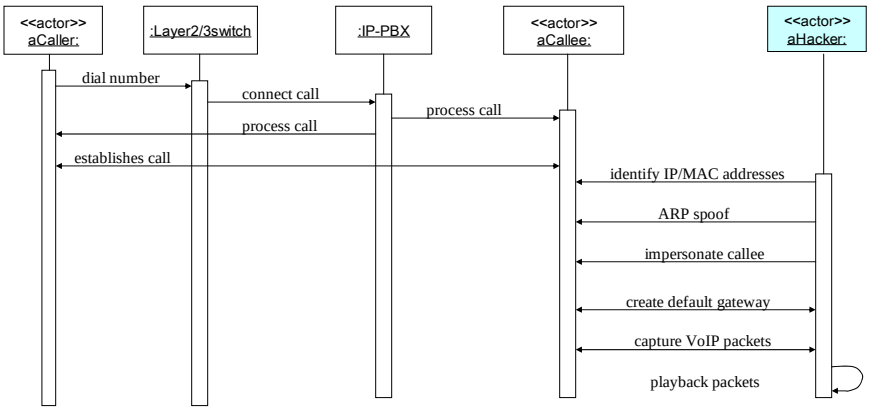

Figure 7 Sequence diagram for a call interception

Likewise, the FragRouter tool would have to be enabled on the attacking machine so the data packets would reach their ultimate destination. The tools used for this purpose can be downloaded freely on the internet.

Also, if the hacker has access to the local switched segment, he may be able to intercept a call by inserting a phone into the voice segment with a spoofed Media Access Control (MAC) address, and assuming the target phone's identity.

A hacker breaking into a VoIP data stream has access to many more calls than he would with traditional telephone tapping. Consequently, he has a much greater opportunity of obtaining useful information from tapping a VoIP data stream than from monitoring traditional phone systems.

The risk of experiencing Call Interception is somewhat limited because it would require physical access to the local network or remote access to a compromised host on the local network. Intercepting voice traffic as it crosses the Internet is more difficult because once the packetized voice hits the carrier, it becomes much harder to single out among other traffic.

Repudiation attacks can take place when two parties talk over the phone and later on one party denies that the conversation occurred. This type of attack can be easily mitigated with a good user authentication as it will be explained in the next section.

Call Hijacking or Redirect attacks could replace a voice mail address with a hacker-specified IP address, opening a channel to the hacker [Gre04]. In this way, all calls placed over the VoIP network will fail to reach the end user. The tools used to launch this kind of attack are similar to those used in call interception.

Denial-of-service (DoS) attacks prevent legitimate users of a network from accessing the features and services provided by the network. One method to launch this type of attack is to flood the server with repeated, requests for legal service in an attempt to overload it. This will cause severe degradation or complete unavailability of the service.

A flooding attack can also be launched against IP phones and Gateways in an attempt to interrupt communications. Often the Ping command is used to carry out such flooding attacks. Ping uses ICMP (Internet Control Message Protocol) [Fer05a]. Attackers can also use the TCP SYN Flood attack to obtain similar results. Since these kinds of attacks can be originated from a wide variety of persons and locations, they are very difficult to mitigate.

Similarly, out-of-sequence voice packets (such as receiving media packets before a session is accepted) or an excessively long phone number could open the way to buffer overflows. VoIP spam might paralyze a number through repeated calling [Gre04]. Thus, unless these DoS attacks are effectively countered, they could make a voice segment unusable.

Signal protocol tampering occurs when a malicious user can monitor and capture the packets that set up the call. By doing so, that user could manipulate fields in the data stream and make VoIP calls without using a VoIP phone [Pog03]. The malicious user could also make an expensive call, and mislead the IP-PBX into believing that it was originated from another user.

Attacks against Softphones occur because as they reside in the data VLAN, they require open access to the voice VLAN in order to access call control, place calls to IP phones, and leave voice messages. Therefore, the deployment of Softphones provides a path for attacks against the voice VLAN.

VoIP systems are capable of handling large volumes of calls using both IP phones and Softphones. Unlike traditional phones, which must be hardwired to a specific PBX port, IP phones can be plugged into any Ethernet jack and assigned an IP address. These features not only represent advantages but also they may make them targets of security attacks.

As mentioned before, Softphones are PC's enabled with voice capabilities, therefore they are especially susceptible to attacks. Softphones are less resilient under attack than IP phones. PC-based IP Phone hosts are more susceptible to attacks due to the number of vulnerabilities of the PC itself (OS and application vulnerabilities).

Note that all these attacks apply also to conference calls and some may apply to the use of voice mail.

4.2.2 Registration attacks

Brute Force attacks are simply an attempt to try all possible values when attempting to authenticate with a system or crack the crypto key used to create ciphertext [Bre99]. For example, an attacker may attempt to brute-force attack a Telnet login, he must first obtain the Telnet prompt on a system. When connection is made to the Telnet port, the hacker will try every potential word or phrase to come up with a possible password. Hackers may attempt in this way to gain unauthorized access to the voice services. These kinds of attacks can be initiated on both the outside and inside of a network constraint.

Reflection attacks are specifically aimed at SIP systems. It may happen when using http digest authentication (i.e. challenge-response with a shared secret) for both request and response. If the same shared secret is used in both directions, an attacker can obtain credentials by reflecting a challenge in a response back in request. This attack can be eliminated by using different shared secrets in each direction. This kind of attack is not a problem when PGP is used for authentication [Mar01].

The **IP Spoofing** attacks described earlier can also be classified as registration attacks.

5. VoIP Security architecture

Security in VoIP can be implemented in two different ways: by using built-in security mechanisms of VoIP protocols (e.g. H.235); or through existing network security standards (e.g. TLS and IPSec).

5.1 Security Mechanisms to counter attacks against VoIP calls

5.1.1 Tunneling and Segmentation

The simplest method to counter Call Interception and other related attacks is to route the voice traffic over a private network using either point-to-point connections, a carrier-based IP VPN service, or a frame relay network running over an ATM core [Con02].

Tunneling is the encapsulation of data from one protocol into the protocol stream of another. Tunnels are by topology "point to point" virtual connections between a network ingress point and a network egress point. At the ingress point, data is encapsulated using encryption, while at the egress point, data is de-encapsulated into the original source format.

Tunnels can be established across a private Metropolitan Area Network (MAN) or a Wide Area Network (WAN), such as Optical Ethernet, Frame Relay, Leased Line, etc., or across the public Internet. In this way, most VoIP traffic, except for virtual private networks (VPN), will not traverse a public network such as the Internet. Tunneling over WANs eliminates the risk of exposing a network to intruders, which comes with opening ports on a firewall to allow VoIP to flow through. Therefore tunnels provide secure transport of the VoIP traffic over the public network.

A secure voice network which employs end-to-end encryption will introduce more latency. This is the major disadvantage in this approach. Latency is a very important issue as users will not accept long call setup times, delay in conversation, or choppy voice quality (jitter). Participants in a conversation will begin to talk over each other as latency increases above 200ms.

VPNs are cost-effective because users can connect to the Internet locally and tunnel back to connect to corporate resources. This not only reduces overhead costs associated with traditional remote access methods, but also improves flexibility and scalability, at the expense of some security.

Network Segmentation consists on keeping voice and data on separate VLANs to segregate VoIP from data traffic. The use of Virtual LANs (VLAN) will provide some quality of service benefits as well as add another layer of complexity for an attacker trying to "sniff" or capture packets off the network. Segmentation is a good idea for increasing security and performance [Pog03]. With good network segmentation, an attack aimed at the data network won't impact critical voice traffic.

Combining data and voice segmentation is another solution for confidentiality attacks as the switched infrastructure mitigates attacks like call interception. To some extent, keeping the segments separate prevents devices in the data VLAN from listening to calls in the voice VLAN. The segmentation technique can be implemented using as a minimum a packet filter on the routing device that provides connectivity between voice and data VLANs.

Additionally, segmentation is important to counter attacks against the voice VLAN when using Softphones which reside in the data VLAN, and for their operation, they need to use the voice VLAN. As mentioned in previous sections, both SIP and H.323 protocols use UDP for media packets. UDP requires wide port ranges in order to allow pinpoint access between the VLANs, thus creating vulnerabilities in the use of Softphones.

Figure 8 illustrates a segmentation technique in VoIP that is achieved by sending voice and data on separate VLANs. A stateful firewall is used in the data VLAN in order to prevent attacks against the voice VLAN when using softphones. On the other hand, the voice VLAN uses a proxy firewall to solve the firewall/NAT traversal issue.

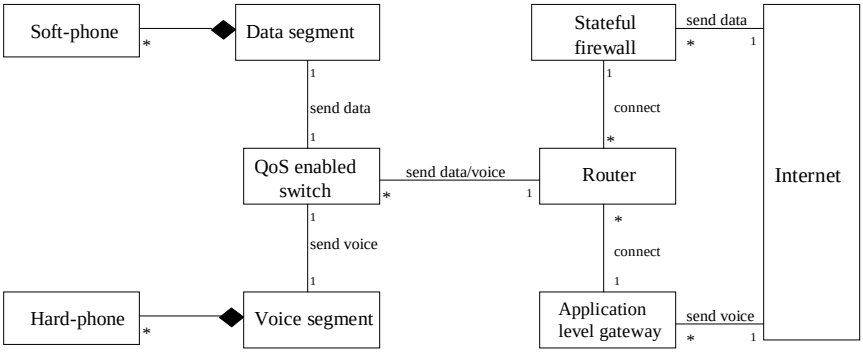

Figure 8 VoIP Segmentation

5.2.1 Encryption

Voice packet encryption is the best defense against call interception. On the other hand, encryption can introduce delay to voice packets and adversely affect the performance in VoIP networks. To improve efficiency and call quality, some devices use hardware cryptosystems rather than those performed in software to reduce the amount of processing time for encryption/decryption. In this way, end devices (e.g.

IP phones) or at least gateways, can be equipped with the required computational power to establish an end-to-end security where messages are encrypted or/and authenticated along the entire route from the caller to the callee.

VoIP designers can also perform encryption at routers. By this method, voice packets are sent from a phone or VoIP gateway to an encrypting router. From there, they are sent to a remote encrypting server which decrypts the message and transmits it to the receiving phone. However, there is a problem, as the data is vulnerable to interception between the phone and the encrypting router on both ends of the connection. This means that the use of this scheme exposes the voice data stream to internal attacks.

It is also possible to perform encryption only at the link-level. Even though asymmetric methods of encryption require high processing power, gateway devices are normally designed to process heavier loads and this method should be transparent to the users. Encryption could possibly be limited to specific fields within the VoIP packets, containing sensitive information.

The class diagram shown in Figure 9**.** shows a Secure-channel communication in VoIP (adapted from the Cryptographic pattern in [Fer05a]).This diagram generalizes the cryptographic transformation in a VoIP call and distinguishes the Caller and Callee roles from those performing the encryption (i.e. Encrypter and Decrypter). Participants in the voice call must agree previously on the data encryption standard (i.e. DES, 3DES, AES) and on a shared secret key.

5.3. Security mechanisms to counter Registration attacks

5.3.1 Authentication

Authentication is the best countermeasure against Registration attacks. In general, authentication is based either on using a shared secret or on public key based methods with certificates. The Diffie-Hellman key exchange is an alternative to prior establishment of a shared secret or password. This method is based on generating or exchanging the shared secret key using public key cryptography (PKI).

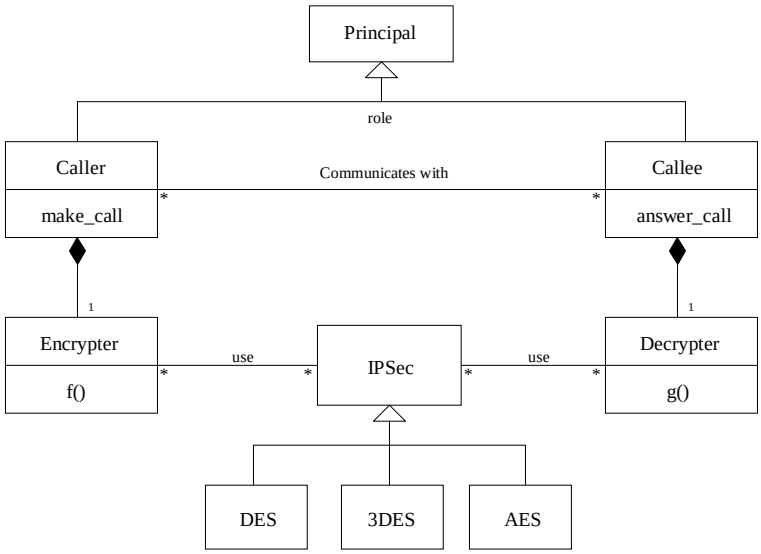

Figure 9 Class Diagram for a VoIP Secure Channel

It is necessary for the authentication mechanisms in a VoIP environment to be selected according to the roles and rights mentioned in the previous section. These mechanisms must reflect the authorization level into the network. Public key cryptography based authentication is the only means of authentication that scales up to arbitrarily large networks by making it possible to securely distribute keys relatively easily through unsecured networks [Mar01].

By using Public Key infrastructure in VoIP, an enterprise-network-based call between two phones (either IP or standard) can be established after it is first encrypted, using the caller's private key and the public key of the remote user. User authentication along with call logging is necessary in order to prevent repudiation in VoIP. Call logging is provided by the IP-PBX. Additionally the IP-PBX will prevent unknown terminal devices from being configured protecting the VoIP system from theft of service.

Authentication in a VoIP network is used to effectively counter attacks in which a hacker spoofs a MAC address and attempts to assume the identity of its target. *Device* authentication is done with the use of MAC address security with network IDSs enabled on closet switches to notify the network administrator about unknown devices. The use of MAC security will greatly reduce the risk of a user (either malicious or in error) connecting an unauthorized device to the voice VLAN and receiving or denying service to other users. Also firewalls that perform source route tracing can be configured to divert IP spoofing. A strong authentication is also the best countermeasure for theft of service attacks.

Likewise, the H.235 component of H.323 specifies two types of authentication [Wei01]:

- *Symmetric***,** is a method of authentication that is less processor intensive and requires no previous communication between the two devices.
- *Subscription based***,** is a method which can be either symmetric or asymmetric. It requires the sharing of a secret key or certificate before the communication can occur. Asymmetric encryption methods are generally very secure; however, they require large amounts of CPU processing power and time.

Subscription-based authentication has three variations, which are:

- password-based with symmetric encryption;
- password-based with hashing (also symmetric);
- certificate-based with signatures (asymmetric).

Authentication is a computationally intensive operation that benefits from the same hardware acceleration as encryption and placing authentication in the phone creates better end-to-end protection than having it done in a router or server.

As previously mentioned, there are other authentication mechanisms that can be specified as those to be used with SIP, they are IPSec based connection and TLS. IPSec uses either The Authentication Header (AH) or The Encapsulating Security Protocol (ESP) for providing cryptographic authentication to IP (v4 and v6) datagrams. The authentication data is computed by using any of the standard message digest algorithms such as HMAC-MD5 and HMAC-SHA. IPSec also uses the Internet Key Exchange (IKE), for establishing shared and authenticated secret keys. This mechanism has to be employed when using IPSec to secure large-scale VoIP networks [Sta02].

5.3.2 Threshold-based analysis

Comparing traffic patterns against predefined thresholds is a simple and effective method of detecting theft-of-service attacks. This system is based on the knowledge that most losses to service providers are caused by large-scale commercial fraud. Such a method can produce an alert, for example, the moment the number of calls being made from a certain location exceeds the threshold of calls defined for that location. This method can be used to successfully recognize and contain theft of long, short, and/or expensive calls.

The straightforward nature of this algorithm allows simple, efficient implementation, thus supporting the large amount of traffic carried over VoIP networks. It does, however, require fine-tuning with respect to the actual setting of thresholds, as the latter must be performed meticulously for each customer and point of contact. Moreover, this technique does not detect several types of fraud [IEC01].

6. Conclusions

This paper presents various ways in which VoIP system designers can more effectively streamline the use of IP telephony as a secure and convenient method of safeguarding and protecting voice communication. We have discussed existing VoIP architectures and provided UML models for their architecture. This provides a precise framework where to apply security. We also considered possible security attacks and related them to the ways the system is used. This is a convenient and systematic way of finding most attacks. Finally we considered some defense mechanisms and provided UML models for some of them. These solutions can be expressed as patterns.

In conclusion, the best security approach in VoIP is to encrypt all voice traffic and to use VPNs to separate VoIP from data traffic in order to increase security and performance; even though it may not be appropriate for all environments. This would ensure that the critical voice traffic would be unaffected if an attack did occur on the data network. Security on VoIP networks can be better implemented using

filtering and/or firewalls to control the traffic between the voice and data VPN. In the near future, if you make a telephone call, it is very likely that it will be over the Internet or some other packet network.

Acknowledgements

This work was supported through a Federal Earmark grant from DISA, administrated by Pragmatics, Inc.

References

[Bas04] Bastermagian N., "Including VoIP over WLAN in a Seamless Next-Generation Wireless Environment", Texas Instruments, March 8, 2004

[Bre99] Brenton, Chris, "Mastering Network Security," Network Press, San Francisco, 1999

[Con02] Conry-Murray, Andrew, "Emerging Technology: Security and Voice over IP- Let's talk", November 4, 2002, http://www.networkmagazine.com/article/NMG20021104S0004

[DSQ01] DSQ Software Ltd. "Voice over Internet Protocols". February 14, 2001 http://www.dsqsoft.com/library/articles/voip_pap.pdf

[Fer04] E.B.Fernandez, "A methodology for secure software design", 2004 Intl. Symposium on Web Services and Applications (ISWS'04), Las Vegas, NV, June 21-24, 2004.

[Fer05a] E. B. Fernandez, E. Gudes, and M. Olivier, Secure Software Systems, Addison-Wesley 2005 (to appear).

[Fer05b] E.B.Fernandez, M. M. Larrondo-Petrie, N. Seliya, N. Delessy-Gassant, and M. Schumacher, "A pattern language for firewalls", to appear in M. Schumacher, E.B.Fernandez, D. Hybertson, F. Buschmann, and P. Sommerlad (Eds.), Security Patterns, Wiley 2005.

[Gre04] Greenfield, David, "Securing The IP Telephony Perimeter", April 5, 2004 http://www.networkmagazine.com/shared/article/showArticle.jhtml?articleId=18900070

[IEC01] The International Engineering Consortium, "Fraud Analysis in IP and NGN", April 12, 2001. http://www.iec.org

[Jai00] Jain, Raj, "Voice over IP: Protocols and Standards," February 7, 2000 http://www.cse.ohio-state.edu/~jain

[Mar01] Marjalaakso, Mika. "Security requirements and Constraints of VoIP", September 17 2001 http://www.hut.fi/~mmarjala/voip

[Mil02] Miller, Mark. Voice over IP Technologies, M & T Books, New York, 2002

[Pog03] Pogar, Noel, "Data Security in a Converged Network" July 23, 2003 http://www.siemens.com/

[Sta02] Stallings, William. "Network Security Essentials: Applications and standards", Prentice Hall, Upper Saddle River, 2002, 5 – 21

[Tan03] Tanenbaum, Andrew, Computer Networks, Prentice Hall PTR 2003

[Wei01] Weiss, Eric, "Security concerns with VoIP" August 20, 2001 http://www.sans.org/rr/papers/index.php?id=323

Biographic Information

Juan C PELAEZ. Mr. Pelaez is a PhD student in the Department of Computer Science and Engineering of Florida Atlantic University. He is part of the Secure Systems Research Group. He is also a Teacher Assistant and Tutor for the Division of Engineering Student Services at FAU.

Eduardo B. FERNANDEZ (http://polaris.cse.fau.edu/~ed), is a professor in the Department of Computer Science and Engineering at Florida Atlantic University, and the leader of the Secure Systems Research Group (http://www.cse.fau.edu/~security). He has published numerous papers and three books on different aspects of security, object-oriented analysis and design, and fault-tolerant systems. He holds a Ph.D. degree from UCLA. His industrial experience includes 8 years with IBM and consulting with several companies.

Authorization and Disclaimer