Voice over IP Security
A layered approach

Amarandei-Stavila Mihai

xmco | Partners

Amarandei-Stavila Mihai *mihaia@gmail.com*
Coordinator: Frédéric Charpentier

Table of contents

Who we are ?

XMCO has been built upon proven, hands-on Penetration Testing methodologies used against the systems of the major international firms around the globe.
We believe that our consultants must keep their hands dirty in the field by consulting as we believe that an equal emphasis on theoretical and real world experience is essential for effective results.
Our pentests naturally include the latest vulnerabilities and defenses. XMCO consultants are willing to share their business skills needed to identify protection opportunities, justify testing activities and optimize security controls appropriate to the business needs in order to reduce business risk.

We go far beyond simply "hacking" your system, which is currently proposed by our competitors. We help our clients to reach the level of security their systems deserve.

More :
info@xmcopartners.com
http://www.xmcopartners.com

1 Voice over IP Security

1.1 A General Overview of Voice over IP

Integration of communication services into the IP network infrastructure, and the Internet especially, is natural course that was started long ago with e-mail, continued with instant messaging and now taken one-step further with integration of standard, classical services like telephony. Voice over IP – the transmission of voice over packet-switched IP networks – is one of the most important emerging trends in telecommunications. Two factors have an important role in the growing popularity of Voice over IP networks: the cost savings factor, inherent to the migration from standard to Voice over IP networks, and the flexibility factor that allows new services and new applications to be added to standard telephony services (video transmission, conferences, etc). With 2005 declared as the Voice over IP (VoIP) year and with predictions of fairly large budgets attributed to VoIP projects in the near future, this technology seems set to replace classic, circuit-based telephony in the near future.

Even if they serve the same purpose, VoIP has a very different architecture from classic telephony. In VoIP networks voice and signalling are multiplexed and travel as normal data inside LANs, WANs or the Internet whereas in classical telephony each conversation has a private, physical, circuit and a dedicated infrastructure that serves only for its transmission. VoIP sound is sampled, quantified, encoded with an appropriate codec and streamed over traditional network architectures. It is and it behaves as normal IP data but at the same time has to obey to the rules imposed by classical telephony in terms of quality of service and availability. Developing a robust architecture that respects all these constraints and is secure is not an easy task, and the fact that many companies have implemented and tried to impose proprietary architectures has added a factor of uncertainty to the strength of this new technology. In the last period however, major companies and institutions have joined in a common effort to create a basic robust standard for VoIP architectures, and security beneficed from a special emphasis with the creation of such projects as VoIPSA

As with many new technologies, VoIP introduces new security risks and new opportunities for attack. Inheriting from both networks and telephony, VoIP is subject to security issues coming from both areas. Classical telephony security issues involving signalling protocol manipulations, phreaking (see [4] for more details) as it was dubbed in the seventies, find their mirror in VoIP specific protocol manipulations. The main purpose remains the same: fraud. Network security issues on the counterpart are far more complex and offer larger perspectives of attack than traditional phreaking. From physical layer to faulty applications, all network security items are relevant to VoIP security. In terms of exposure, the transport of voice data over the Internet, a highly insecure and unreliable environment, multiplies the attack surface and will surely lead to more attacks on this technology. Furthermore, the synergies of these two aspects of VoIP emerge to add new security threats such as signalling protocols Denial of Service.

1.2 Voice over IP Particularities

One of the common mistakes when approaching Voice over IP Security is to treat this technology and its applications like all other network-based applications and therefore use classical security measures. This approach fails to consider the fundamental property that makes VoIP different from classical network-aware technologies: Voice over IP is time critical. It demands Quality of Service (QoS) and is generally deployed over an environment where QoS does not exist as a standard feature.

Voice over IP is a highly demanding technology. It is time critical and therefore requires a mechanism for assuring the Quality of Service. The bounds of the quality factor are not very permissive, with latency limit at 150ms and packet loss limit at 3%. By adding all this issues with other unquantifiable disrupting factors such as jitter, we can obtain a first image of the greatest weakness of VoIP, its high sensitivity to disruptions and therefore to disruptive attacks.

These type of attacks are popularly known as Denial of Service (DoS) attacks. VOIP-specific attacks (i.e., floods of specially crafted signalling messages) can and will result in DoS for many VOIP-aware devices and applications. SIP hardware phone endpoints may freeze and crash when attempting to process a high rate of packet traffic. SIP proxy servers may experience failure and intermittent log discrepancies with a VOIP-specific signalling attack of under 1Mb/sec. In general, the packet rate of the attack may have more impact than the bandwidth; i.e., a high packet rate may result in a denial of service even if the bandwidth consumed is low. This makes DoS attack easier to conduct against unprotected VoIP networks. If with normal DoS high bandwidth is often required to “choke off” the target attacked, with VoIP the target can “choke” on the amount of data it has to process and the fact that the time spent processing this unnecessary data degrades VoIP QoS performances below the acceptable levels.

1.3 Voice over IP Architectures

In a simplistic view, The Voice over IP technology has the goal of establishing and managing communication sessions for transmitting voice data, or sound in general, over standard IP networks. VoIP may additionally support the transmissions of other data formats such as video, text or images. A stable and reliable transmission has to be maintained all throughout the conversation, and the session needs to be ended when either of the pairs with the needed authority decides of this.

Two types of protocols are used by VoIP technology, in a similar manner to standard telephony: signalling protocols and media transport protocols.

Today, two VoIP architectures have emerged as standards and are widely deployed and used throughout the world. Their names reflect the signalling protocol they use. They are H.323 and SIP architecture.

H.323 is the International Telecommunication Unit (ITU) standard for audio and video transmissions over packet based networks. H.323 is actually a wrapper standard, encompassing several other protocols, including H.225, H.245, and others. H.323 uses RTP as standard protocol for media transport. Protocols appertaining to the H.323 standard are binary protocols, based on the ASN.1 standard[1]. Each of these protocols has a specific role in the call setup process, and all but one are made to dynamic ports. An H.323 network is made up of several endpoints (terminals), a gateway, and possibly a gatekeeper (address resolution and bandwidth control), Multipoint

[1] **Abstract Syntax Notation one** : A standard and flexible notation that describes data structures for representing, encoding, transmitting, and decoding data.

control unit (for multi point conferencing), and Back End Service (for storing and maintaining configuration data about endpoints).

SIP is the Internet Engineering Task Force (IETF) protocol dedicated to initiating two-way communication sessions. SIP gives its name to the VoIP architecture relying on it for signalling functionality. It is important to emphasize that SIP is not specific to VoIP and can be used in any session driven technology/application. SIP is text based, similar to HTTP. It can be carried by TCP, UDP, or SCTP. UDP may be used to decrease overhead and increase speed and efficiency, or TCP may be used if SSL/TLS2 is incorporated for security services. SIP uses one network port with the default value 5060.

The architecture of a SIP network is different from the H.323 structure. A SIP network is made up of end points (terminals, User Agents - UA), a proxy and/or redirect server (for endpoints' message transmission), location server(for locating users), and registrar(for registering location information). The registrar and location server may be integrated in the proxy server.

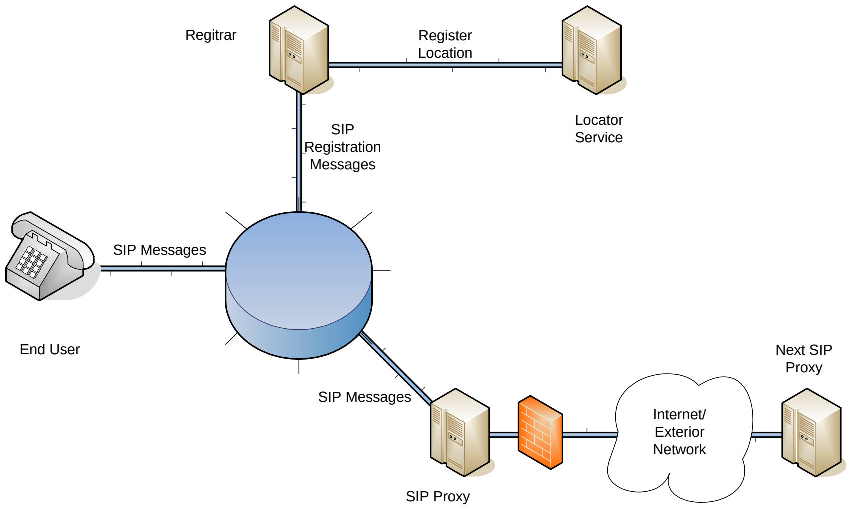

Figure 1.1
SIP Architecture

[2] **Secure Sockets Layer** (**SSL**) and **Transport Layer Security** (**TLS**), its successor, are cryptographic protocols which provide secure communications on the Internet. There are only slight differences between SSL 3.0 and TLS 1.0, but they are not interchangeable

1.4 Voice over IP Threats

At the moment this paper is being written Voice over IP Security makes the headlines of all major specialized media. From *The Register* to *ZdNet* and from *Slashdot* to *eWeek*, all major IT medias seem to be interested in this subject. The reasons are simple: the promises of the this technology and the lack of expertise and standards in VoIP Security. Many security experts have warned that up until now VoIP has been developed and deployed with a purely commercial goal in mind, overlooking security and focusing on functionality. As the technology matures, becomes widely available and usable, security issues gain interest. However, no real standards exist yet. VoIPSA (Voice over IP Security Alliance), of which I am a member, aims to be the first organization building a solid and coherent taxonomy of vulnerabilities and attacks threatening VoIP infrastructures. However, the results are not close. Given the current situation, the following paragraphs depict a personal vision on the possible structure of VoIP potential threats.

The approach I have chosen in organizing and discovering VoIP threats is a layered, source based one. The layered view was suggested by two different reasons. First, partial analogies that can be drawn between the TCP/IP protocol stack and the VoIP SIP protocol architecture as some layers can be considered common to the two concepts, namely physical, application and network. Furthermore, the resemblance between SIP session initiation and TCP three-way handshake has already been underlined in the previous section and is often cited in specialized literature such as [7]. This resemblance creates analog attacks, as we shall see later in this section. Second, a source based layered approach permits to quickly spot out solutions to the potential threats discussed. Restraining the attacks to a specific layer restricts the initial solution research space to that layer and makes addressing the issue easier. Often, problems originating from one lever can be dealt with within that level. This is not always the case as we shall see, but if a security issue will not find a viable solution in its layer, the other layers can be researched one by one for methods of blocking the attack. In the next sections I will concentrate on two specific VoIP layers : signalling protocols and Transmissions protocols. I will also concede some space to the application layer, vital to any IT infrastrcture.

Before defining and detailing these actual layers, it is important to build a small glossary explaining the general attacks foreseeable in VoIP networks. The following definitions and short descriptions are based on [4]:

Denial of Service

A **denial-of-service attack** (also, **DoS attack**) is an attack on a computer system or network that causes a loss of service to users, typically the loss of network connectivity and services by consuming the bandwidth of the victim network or overloading the computational resources of the victim system.

A DoS attack can be perpetrated in a number of ways. There are three basic types of attack:

1. consumption of computational resources, such as bandwidth, disk space, or CPU time
2. disruption of configuration information, such as routing information
3. disruption of physical network components

Eavesdropping

Eavesdropping is the intercepting and reading of messages and conversations by unintended recipients. In VoIP, eavesdropping is an attack giving an attacker the ability to listen and record private phone conversations. Eavesdropping can have important, unexpected consequences as people may use telephone systems for divulging crucial information such as Social Security numbers, Credit Card numbers or any other confidential information. Inside a company, eavesdropping could allow access to confidential business information.

Man in the Middle

Man in the middle attack (**MITM**) is an attack in which an attacker is able to read, insert and modify at will, messages between two parties without either party knowing that the link between them has been compromised. The attacker must be able to observe and intercept messages going between the two victims. MITM attack can be used for conducting other sub attacks such as eavesdropping or DoS

Call hijack

In VoIP and classical telephony, **Call Hijack** attack refers to a situation where one of the intended end points of the conversation is exchanged with the attacker. Call hijacking can have common consequences with eavesdropping attacks (access to confidential information). Generally if an attacker is able to conduct a Call Hijacking attack, he will evolve it into a MITM attack to avoid raising suspicions.

Spoofing attack

A **spoofing attack**, in computer security terms, refers to a situation in which one person or program is able to masquerade successfully as another.

Call Fraud

Call fraud is an attack specific to telephony and VoIP, in consists in the illicit use of a VoIP infrastructure to place phone calls. These phone calls seem to originate from legitimate users inside the attacked network/system.

1.4.1 Signaling Protocols Layer

The attacks falling in this category are partially related to the signaling attacks on the classical phone network that made the headlines in the '70s. We could speak of network phreaking if we wanted, but the situation has dramatically changed.

This layer is the most important and the most VoIP specific attack layer. Several attacks originating from this layer could be carried with minimal resources and could have disastrous consequences. The most important and potent attacks are:

1.4.1.1 SIP based Denials of Service

SIP Bombing

This attack involves transmitting a large quantity of forged SIP messages to a targeted VoIP system. As discussed in earlier sections IP telephony is very sensitive to availability issues. A large quantity of bogus SIP messages will require the allocation of computational resources for decoding and interpreting. As the system is busy treating the bogus messages, valid ones will be treated at a much slower rate and overall performance of VoIP conversation will decay.

This type of attack has already been tested successfully on VoIP devices. Hardware IP phones have crashed on rebooted when subjected to heavy SIP traffic. Softphones seem to be more adapted, but some applications have crashed or become unresponsive and unusable

SIP-Cancel/Bye DoS

Two distinct malicious actions are treated here. They are reunited in the same category because the attack blueprint is the similar for the two cases: an attacker sends a crafted SIP message to a victim user upon detecting that the victim established a SIP session/VoIP conversation with another user. Figures 1.2.a and 1.2.b illustrate this type of DoS.

An effective SIP-CANCEL attack has to be launched in the time window available between the sending of the INVITE method and the receiving of the last ACK that ends the SIP session initiation handshake. This leaves a very restrictive window to the attacker. Considering the small attack window available for a SIP-CANCEL DoS, we can conclude that this attack will most likely be successfully employed from within the victim's local network, where taking advantage of the INVITE-ACK windows is much easier. In this particular situation, an attacker can issue a CANCEL immediately upon detecting that its victim received an INVITE request. How the attacker detects this event is out of this attacks scope.

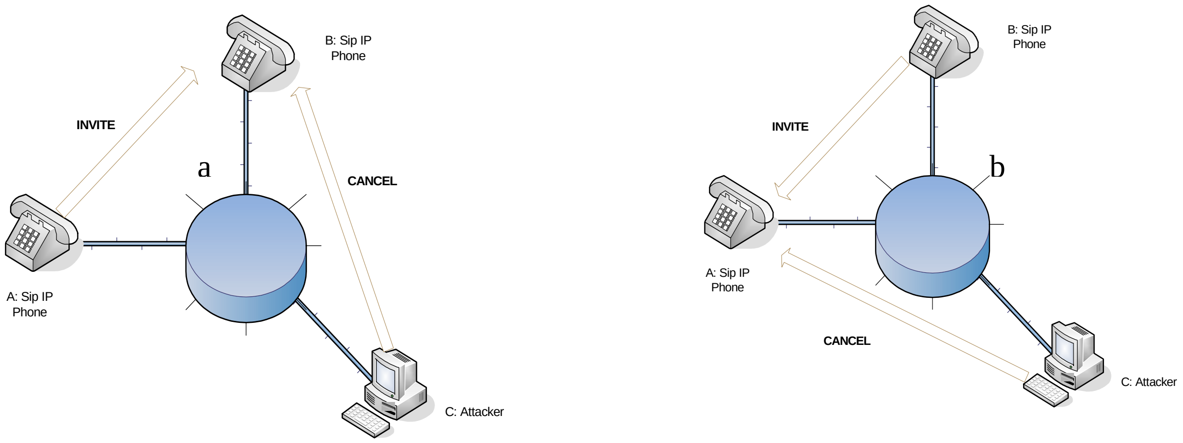

Figure 1.2
SIP Cancel DoS

3.7.a *A is denied making phone calls*
3.7.b *A is denied receiving phone calls*

In opposition with SIP-CANCEL, SIP-BYE, can be launched at any moment after a VoIP conversation has begun.The attack window available is therefore the entire duration conversation, which makes a successful attack much easier to realize.

The two attacks described earlier have a reasonable number of *mitigating factors* that makes it difficult to conduct them in real-live scenarios:

- Attack opportunity detection: these attacks are opportunity based attacks, that is they target special behaviors and situations. Therefore, successful attacking has to synchronize with the presence of the targeted behavior/situation..
- Header reproduction: a successful attack has to reproduce the original headers of the SIP session in the CANCEL/BYE messages for them to be accepted as legitimate messages.

1.4.1.2 SIP based Man in the Middle/Call Hijacking

Since the two attacks have common consequences and common attack blueprints, we will treat them as a sole entity. Once a call is hijacked, it a simple to forward it to the original callee, thus realizing a MITM attack. Two main categories of SIP based MITM attacks can be distinguished: manipulation of Registration Records and “3xx response codes” based attacks.(3xx response codes are responsible for redirection procedures)

<u>Manipulation of the Registration Records</u>

This attack enables an attacker, generally based on the same network as the victim, to receive all the victim’s calls by manipulating the registration associated with the victims SIP URI.

A registrar assesses the identity of a UA. The “From” header of a SIP request can be arbitrarily modified and hence open to malicious registration. Furthermore, the UDP protocol used for registration requests is connectionless and can be easily spoofed. SIP registrars are not required to authenticate the UA requesting a registration, and with trusted intranets, some will certainly not. When authentication is used, it is not strong, only involving use of a MD5 digest of the username, password, and timestamp-based nonce sent in the authentication challenge. Extreme cases of these attacks include exterior attackers that are able to successfully register as internal users to the SIP registrar, but this situation is highly improbable.

<u>3XX response codes</u>

The 3xx SIP response codes class corresponds to redirects and informs the requester that further actions has to be undertaken in order to successfully accomplish the request. The 3xx response codes class of SIP based attacks relies on forged responses. The attack outline is the following:

- *the victim* issues a SIP request (an INVITE request for example)
- *the attacker* sends a *3xx code response* to the initiator. The attacker usurps the identity of either the callee UA or one of the SIP components(proxy, registrar, etc)
- the victim’s SIP client receives the *forged 3xx response* and *redirects* its communication through the attackers system for the rest of its request; the attack is complete.

The simplest attack of this type is the Call Hijacking attack using 301 Moved Permanently response code. In this classical scenario, the attacker sends a 301 response upon detecting that the victim has issued an INVITE Request. The victim in turn contacts the attacker for achieving the SIP connection. Alternatively, for single time attacks, 302 Moved Temporarily response code can be used instead of 301. A graphical illustration of this attack is given in Figure 1.3

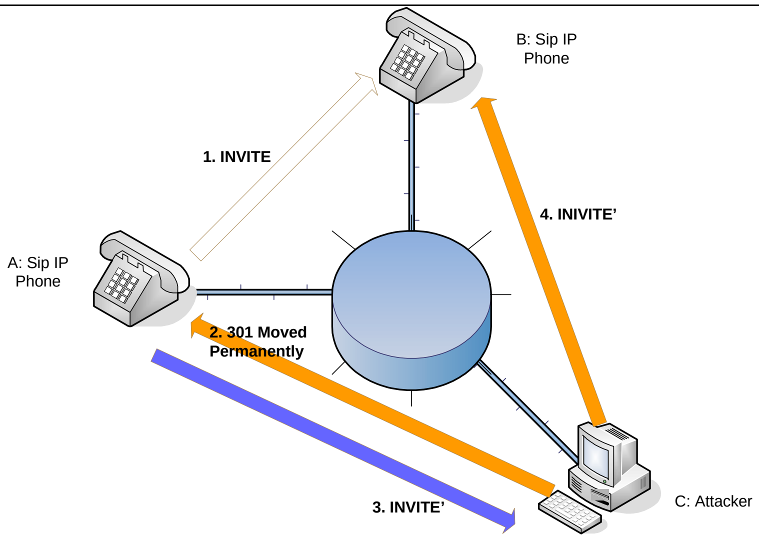

Figure 1.3
Simple 301 Call Hijack

The attacker is not restricted to UAs when choosing his forged identity. By careful manipulation of 3xx responses and by spoofing a proxy server, a variant of the above attack can be achieved with the *305 Use Proxy* response code.

Similar to SIP DoS attacks, *mitigating factors* make it hard to conduct the MITM attacks in real-live scenarios:

- Like SIP based DoS attacks these attacks are *opportunity based* attacks.
- Headers reproduction: in a similar manner to SIP DoS, successful MITM attacks require that the forged responses coming from the attacker machines contains the right header content to be accepted as legitimate. The Cseq field is especially hard to estimate or intercept and thus mirror.
- Authentication mechanisms: Strong authentication mechanisms for the registration process diminish the chances of registration hijacking.

1.4.1.3 Possible solutions for SIP based attacks

All the attacks described in this section rely on tampering with and creating spoofed SIP messages. Whether it is for crafting SIP messages, mirroring headers or simply determining the type of SIP message that a victim has issued, the clear text format of SIP greatly helps the attacker's goal. A standard, clear-text SIP message can easily be forged or replayed if a digital signature or a strong encryption scheme does not protect it.

The solution for avoiding the attacks described earlier is protecting SIP content from tampering, interception, and reply with strong encryption and authentication mechanisms. Two particular solutions are becoming standard:

TLS usage within SIP

SIP's RFC imposes the use of TLS (Transport Layer Security) for proxies, redirect servers, and registrars to protect SIP messages. Using TLS for end points is recommended. TLS is able to protect SIP messages against integrity, confidentiality issues and against replay. It provides integrated key-management with mutual authentication and secure key distribution. TLS is applicable hop-by-hop between UAs/proxies or between proxies. The problem of TLS utilization with SIP is the requirement of a reliable transport protocol (TCP-based SIP signaling). TLS cannot be applied to UDP-based SIP signaling. Just as secure HTTP is specified with the “https:”, secure SIP is specified with a Universal Resource Indicator (URI) that begins with “sips:”.

IPsec usage within SIP

IPsec may be used to provide encryption for SIP messages at the network layer. This type of security is most suited to securing SIP hosts in a SIP VPN scenario (SIP user agents/proxies). IPsec works for all UDP, TCP and SCTP based SIP signaling. IPsec may be used to provide authentication, integrity and confidentiality for the transmitted data and supports end-to-end as well as hop-by-hop encryption. At this time there is no default cipher suite for IPsec defined in SIP. SIP’s rfc does not mandate a particular cipher mechanism to be used with SIP over IPSec scenarios. One accepted protocol for key management is Internet Key Exchange (IKE).

The IKE protocol provides automated cryptographic key exchange and management mechanisms for IPsec. IKE is used to negotiate security associations (SAs) for use with its own key management exchanges (called Phase 1) and for other services such as IPsec (called Phase 2). IKE is particularly used in the establishment of VPNs.

1.4.2 Transport Protocols Layer

In SIP VoIP networks, transport protocol based threats contain attacks taking advantage of weakness of RTP/RTCP protocols. They rely upon a non-encrypted RTP stream. Several categories of threats can be defined.

1.4.2.1 Eavesdropping

The first attack of this category, and the most simple to carry out, is the eavesdrop attack. RTP uses standard codecs to encode audio data it carries. The main weakness of RTP exploited in this case is that information on the used codec is available in the header of every RTP packet, via the PT header field. An attacker with the ability to intercept VoIP media traffic has therefore no problem in saving and later decoding a RTP stream for later auditions.

Demonstrations of such attacks have been made publicly available. Furthermore, tools for conducting this attack are already freely available on the Internet, greatly increasing the potential of this type of attack.

1.4.2.2 RTP Insertion attacks

This class of Transport Level attacks encompasses a large selection of attacks. The common characteristic is that the attacker inserts rogue RTP packets in a RTP stream. Depending on the form of the RTP packets inserted, several outcomes are possible.

SSRC Collision attacks

Upon receiving packets from two different sources with the same SSRC, RTP is put in a collision situation. RTP's collision management mechanism is simple: if a source discovers at any time that another source is using the same SSRC identifier as its own, it must send an RTCP BYE packet (detailed later in this section) for the old identifier and choose another random one; if a receiver discovers that two other sources are colliding, it may keep the packets from one and discard the packets from the other; the two sources are expected to resolve the collision so that the situation doesn't last.

Analyzing how RTP handles collisions, two DoS attacks are foreseeable:

- The attacker "steals" the SSRC of one of the peers and sends its own RTP messages to the other peer. Upon receiving packets with the same SSRC, the receiver chooses to accept packets originating from a single source. The attacker can thus effectively eject a VoIP user from a session.
- The attacker sends to a victim RTP messages labeled with the victim's SSRC. The victim is forced of abandoning current RTP streams in order to choose a new, collision free SSRC. This results in an interruption of any conversation the victim user is involved in.

Other SSRC manipulation

SSRC manipulation can equally be used in a more complex and "innovative" way. If an attacker knows the SSRC of one of the peers, he should be able to forge messages with the same SSRC and IP characteristics, but with higher *timestamp* and *sequence number* RTP values than the legitimate ones. At the receiver end, the RTP application will process the attackers packets first and discard the legitimate packets since they have invalid, older, timestamps. This attack results in the fake content being played before the real audio content. This attack can be seen as a type of Denial of Service.

Codec Manipulation

The RTP protocol supports mid-session codec change. This feature enables RTP to adapt to dynamic network conditions by changing the quality of sound it transports. If the available bandwidth decreases, RTP will detect this like an increase in the number of lost packets and will degrade the codec quality. If more bandwidth becomes available, RTP can do the opposite operation and assure better speech quality.

Attackers can conduct Denial of Service attacks by abusing this feature. An attacker with the ability to insert valid RTP packets in a RTP stream can degrade the conversation quality in two opposite ways. First, the attacker can forge a voice codec change and augment the quality of a RTP stream. This will result in higher bandwidth usage, an increase in packet loss and finally can cripple the conversation up to a point where no audio data can be practically transmitted over the RTP stream. Second, an opposite attack is possible, where the attacker downgrades the codec quality. By forcing a low quality codecs the attacker can degrade the streams' audio quality and again render the conversation impossible. Variations of these attacks involve rapid alternations of codec change that can crash end-systems or force them to consume important resources in this process and induce latency in the VoIP conversations.

1.4.2.3 RTCP insertion attacks

RTCP is the control protocol associated with RTP. RTCP is based on the periodic transmission of control packets to all participants in the session, using the same distribution mechanism and route as the RTP packets. RTCP uses different port numbers than the associated RTP stream. The most important function of RTCP is to analyze traffic conditions and provide feedback on the quality of the data distribution. This is an integral part of the RTP's role as a transport protocol and is related to the flow and congestion control functions of other transport protocols. Based on RTCP parameters such as packet-loss, delay and jitter can be determined. Additionally, RTCP provides a session leave mechanism by the use of a special BYE packet

An attacker with the ability of inserting forged RTCP packets in a RTP conversation can forge this reports and force induce disruptive behaviours.

1.4.2.4 Possible solutions for RTP based attacks

As is the case with some of the SIP based attacks, most RTP based attacks (excepted for eavesdropping) rely on forging spoofed RTP/RTCP packets and inserting them in the RTP stream. Without protection, specialists consider RTP as an insecure protocol. If an attacker has the ability to intercept RTP packets, forging his/her malicious packets and emitting them at the right moments is trivial. Even if the attacker does not have access to the RTP stream, creating rogue RTP packets that will appear as legitimate and thus launching attacks is not a difficult task, given the attacker has some information on the peers involved in the RTP stream. If no such information is available to the attacker, the use of UDP as the transport protocol for RTP makes brute force attacks on RTP unprotected packets' parameters virtually untraceable and easily conductible.

As was the case with SIP, the solution for protecting RTP streams against the attacks described earlier is introducing an encryption mechanism and digital signatures such as secured hashes.

SRTP or Secure Real-time Protocol is the standard for secure Real-time Transport Protocol.

Secure Real-time Protocol (SRTP)

The Secure Real-time Protocol is a profile of the Real-time Transport Protocol (RTP) offering not only confidentiality, message authentication, and replay protection for the RTP and RTCP traffic. SRTP was being standardized at the IETF in the AVT working group. It was released as RFC 3711 in March 2004.

1.4.3 Application Layer

Voice over IP SIP based architecture is just as the name states, an architecture, a style and method of design and construction. Passing from the design state, where we deal with concepts and modelled interactions to practical Voice over IP involves implementations. All the ideas and concepts described all the protocols and the security measures suggested have to be implemented in real software and hardware devices that must function according to the guidelines defined by the architecture for obtaining working Voice over IP networks.

A practical result in this field comes from the PROTOS Test-Suite: c07-sip[3]. PROTOS is a project developed by researchers from University of Oulu, Findland. Its goal is provide a framework for testing known protocol implementations. For each tested protocol, the project releases a Test-Suite that includes the source code together with the tests overview and results. Tests are conducted protocols by injecting of exceptional elements (overflows, unexpected characters, etc) in specific products implementing the tested protocol. The vendors remain anonymous for protecting against attacks issued from the vulnerabilities discovered. The SIP implementations test results were released on the 17 of December 2003. The actual results were disastrous. Out of 9 unnamed products test, none passed all the security tests. Detailed results of these tests can be found at [2].

No wonder solutions exist for protecting against application layer attacks in VoIP networks. The situation is no different from any applicative infrastructure. A good remedy for avoiding attacks based on known software vulnerabilities is an active security watch and a consistent patch management policy. Other partial solutions, based upon security applications such as firewalls and Intrusion Detection Systems, may help protect against some attacks, but the applicative security of a VoIP network should not be solely based on these kind of applications

[3] http://www.ee.oulu.fi/research/ouspg/protos/testing/c07/sip/#h-ref33

1.5 Future threats to Voice over IP Security

Three software properties can be held reliable for the ever rising number of vulnerabilities affecting software. Those three properties, complexity, extensibility and connectivity, are not restrained to software, they can be generalized to all the IT domain and since Voice over IP is part of this domain, it follows the same guidelines.

Complexity in VoIP will first come in the form of intelligent terminals. Already, vendors like Cisco are developing IP Phones capable of running small applications like calendars, agendas but also live results from the stock exchange. As IP Phones become available for software developers, the same process that lead to the intelligent GSM terminals will be applied here. This increase in features and complexity comes however with a security cost: more applications equal more avenues of attack. The programs that will execute on VoIP devices can be affected by vulnerabilities, and if they have additional network connectivity (the stock exchange watch) they can be accessed and manipulated by attackers. We have seen that GSM phones are becoming targets for viruses. I think the question is not whether we will see IP Phones plagued by a virus, but rather how soon?

Extensibility for VoIP devices will likely follow the GSM model as well. Integration of Java Virtual Machines or other products capable of executing mobile code into such devices seems will multiply the number of applications that such a device can execute. This in turn will multiply the number of attack points a hacker may search to exploit and therefore affect the overall security of the VoIP network.

Connectivity is already inherent to VoIP, as network communication is inherent to this technology. A new direction emerging in today’s network model is wireless networks. As users require and enjoy more mobility, it is likely that this tendency will eventually affect VoIP. Wireless networks change the rules of security, as they make impossible to bind a network to fixed, known limits. As VoIP devices start supporting wireless connectivity, VoIP networks will gain a broader exposure and new security mechanisms must be set up to protect against attacks coming from “the air”.

2 Conclusions

The Voice over IP security study had two independent goals: a commercial, product oriented one and a personal, educational one. The commercial purpose of this study is use the taxonomy and threat structure defined for future Voice over IP security audits (the first will be conducted in July 2005). To attain this first goal I had to organize the information available and my research results on VoIP security in a structure adapted for audit usage. For the second, educational goal, the information acquired by security watch, being VoIP specific or simple security information helped me better shape and explain the threats and attack points that could affect VoIP networks. The taxonomy build changed in turn the vision I had on Security Watch and the impact of vulnerabilities.

After conducting this VoIP Security Study we can conclude that VoIP security is in an incipient phase at the moment. Threats and attacks can be defined and theorised, but are difficult to carry out in practice, mainly due to the lack of knowledge and testing opportunities for attackers. However, as soon as VoIP networks will gain more popularity, attackers will probably become very interested in this new technology. The layered approach chosen for the VoIP threats analysis shapes an audit guideline, underlines solutions for combating VoIP threats and has helped me define new attack patterns.

Bibliography

[1] Greg Hoglund and Gary McGraw ***– Exploiting Software : How to break code***. Addison-Wesley, August 2004

[2] John Viega and Gary McGraw – ***Building Secure Software: How to Avoid Security Problems the Right Way***. Addison-Wesley, September 2002

[3] Yves Younan – ***An overview of common programming security vulnerabilities and possible solutions***. Vrije Universiteit Brussel, Agust 2003

[4] Wikipedia Encyclopaedia – ***http://en.wikipedia.org***, Wikipedia, the free encyclopedia

[5] Andrew S. Tanenbaum – ***Computer Networks****.* Byblos Bucharest, 2004

[6] Andrew S. Tanenbaum – ***Modern Operating Systems****.* Byblos Bucharest, 2004

[7] D. Richard Kuhn, Thomas J. Walsh, Steffen Fries – ***Security Considerations for Voice Over IP Systems***. National Institute of Standards and Technology, January 2005

[8] Dorgham Sisalem and Jiri Kuthan – ***Understanding SIP***. Whitepaper, Mobile Integrated Services

[9] J. Halpern – IP ***Telephony Security in Depth***. White Paper, Cysco Systems, 2002.

[10] Networking Working Group – ***RFC2543 – SIP: Session Initiation Protocol***, The Internet Society, 1999

[11] Networking Working Group – ***RFC1889 – RTP: A Transport Protocol for Real Time Applications***, The Internet Society, 1996

[12] Bruce Schneier – ***Attack Trees.*** Dr. Dobb's Journal December 1999

[13] Bruce Schneier – **Schneier on Security Weblog**. http://www.schneier.com/blog/