Voice-over-IP Security Mechanisms – State-of-the-art, risks assessment, concepts and recommendations

Prof. Dr. Evren Eren
University of Applied Sciences Dortmund
Emil-Figge-Str. 42, D-44227 Dortmund
e-mail: eren@fh-dortmund.de, phone: +49-231-755-6776, fax: +49-231-755-6776

Dr. Kai-Oliver Detken
DECOIT GmbH
Fahrenheitstraße 9, D-28359 Bremen
e-mail: detken@decoit.de, phone: +49-421-596064-01, fax: +49-421-596064-09

Abstract

These days most companies and organisations with a distributed structure find Voice-over-IP (VoIP) is a useful technology. However, most of them do not sufficiently reflect upon security issues or merely know about the risks. Therefore, networks, systems and applications are at risk. The rapid growth of the VoIP market and effortless adoption to enterprise IT-infrastructures implicate an increasing need for long-term sustainable security concepts and solutions.

This paper addresses the issues described above in five steps:

- Identifying typical VoIP communication and application profiles in enterprises
- Analysing security risks and possible attacks and their implications on the overall security
- Assessing risks
- Presenting security mechanisms and standards
- Presenting security concepts and recommendations

1. Introduction

Back in 1998 VoIP-technology did not manage to convince CEOs and solutions did not match the requirements of the big market. However, this changed in 2004. It took a long time to introduce VoIP to enterprises and convince CEOs that this technology is worth integrating into existing infrastructures, as it has positive economical (short return-on-invest cycles) and administrative (IP integration and *one* network for data and telephony) effects and benefits.

2. VoIP scenarios and protocols

First of all, it is essential to classify typical VoIP-based communication and application profiles in enterprises and analyse, which VoIP protocols are typically being used.

In the market there are many VoIP protocols (e.g. SIP, RTP, RTCP, SRTP, ZRTP, H.323, MINET, IAX (ASTERISK), MGCP, MEGACO, SCCP (CISCO)), some of them are proprietary and others are open standards. Most popular of the latter are SIP and H.323.

As the implementation of VoIP continues to grow, many enterprises and public organisations deploy hybrid VoIP implementations, i.e. both, circuit-switched and VoIP telephony systems.

Moreover, possible VoIP deployment scenarios are the following:

- **Campus VoIP:** Campus VoIP uses an IP PBX (Private Branch eXchange), which is most common, or IP-enabled PBX. IP phones and/or softphones are connected to the IP PBX. Calls initiated from these phones are routed through a gateway to the PSTN. Thus, this topology is not prone to attacks since the VoIP network does not extend to the Internet or any other non-trusted network. Potential attacks must originate from within the intranet.
- **IP Centrex/Hosted IP:** This type requires the involvement of a VoIP service provider hosting the IP PBX and providing VoIP services from this network. The enterprise only needs IP phones, no other VoIP customer premises equipment is necessary. In this case – as with Campus VoIP – internal attacks exist; additionally, attacks are possible from the service provider's network, since it is a shared one.
- **VoIP Trunks:** VoIP trunks increasingly substitute circuit-switched connections, e.g. T1 and PRI on the basis of non-trusted networks. Especially, attacks coming from the Internet make the enterprise network vulnerable.

Figure 1 illustrates a typical VoIP architecture, which offers great flexibility regarding its scalability for large data networks providing VoIP services. In this architecture, the gateways are located at the edge of the network. Call handling and management has been moved to other devices (call agents/media gateway controllers), which are responsible for voice calls signalling. If one call agent/media gateway controller is out of service for several reasons, another one can take over the responsibility of the gateways. Additionally, the intelligence for call control functionalities has been removed from the hardware (media gateway) to a general-purpose computer.

3. Risks in VoIP-based communications

In this section we discuss risks associated with VoIP-based communications. In order to achieve a holistic approach, we take a multi-level view on the "network layer", "application layer" and the shortcomings of VoIP protocols. In this context, potential threats and attacks are

discussed and compared with each other. Implications on the overall security are discussed.

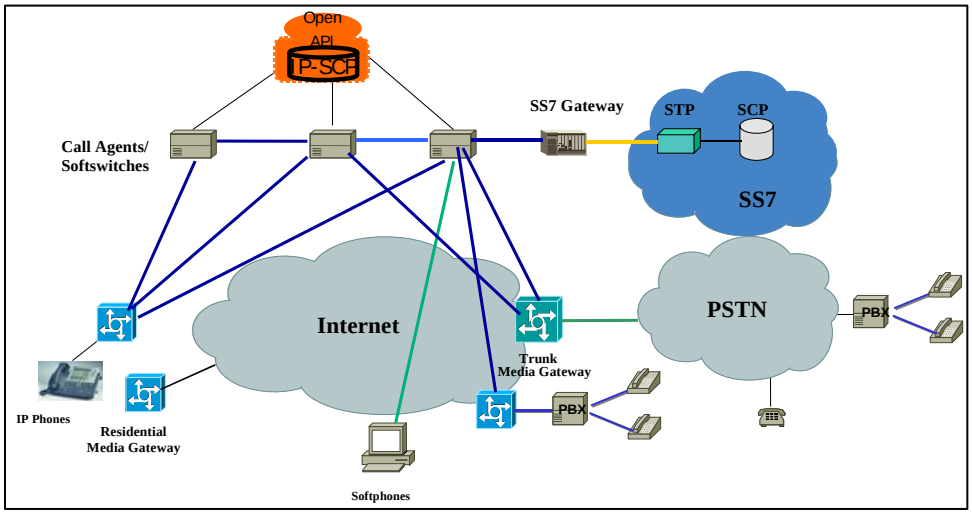

Figure 1: VoIP Architecture

3.1 Shortcomings of the VoIP protocols

SIP (Session Initiation Protocol, RFC-3261) is a plaintext-based protocol operating via UDP or TCP connections. Thus, security vulnerabilities exist as e.g. in SMTP or HTTP. SIP messages are mostly not authenticated and most of the devices do not check the source of the message. Attackers can infiltrate messages to manipulate or disturb SIP services. Also, flooding with connection requests to SIP clients (DoS-attack) is likely. Established connections can be terminated and even directed to unauthorized instances. Furthermore, Directory Harvest attacks endanger SIP. Starting from the domain name of an organisation an attacker tries to seek out valid SIP accounts by calling arbitrary user names.

Typical threats are SIP-Spam (identity forgery), manipulation, redirecting and sniffing of connections, flooding of mailboxes with Spam and modification of messages.

Another category of attacks compromise the registration process. For the registration SIP provides authentication, however, this is optional and some vendors do not implement this feature at all. Also, most registration servers save only little state information of the client. This leads to problems such as unauthorised registration, resource and registration theft, and registration hijacking.

A further class of attack is protocol abuse. A weakness in the protocol architecture allows to modify and to forge the transmitted Caller-ID. A spammer is able to send an arbitrary number of requests, hence forcing the callee to accept the call.

Last but not least, SIP can be abused to transport viruses, worms or trojans, so that SIP applications are as vulnerable as other network applications, because an attack can address both, the SIP application itself and the underlying operating system.

Wrong identities and Man-in-the-Middle (MitM) attacks make the **H.323 protocol suite** assailable. The identification of a caller is managed by an authentication password, which is communicated unencrypted via the network. An attacker simply has to probe the signalling stream, which then can be decoded by an ASN.1-parser, e.g. a packet sniffer like Wireshark (formerly Ethereal). During connection establishment the attacker can modify IP addresses, and possibly redirect the stream to an end-point for sniffing purposes. Both devices and gateways are affected by this threat.

IAX (Inter-Asterisk eXchange; Internet Draft [8]) Version 2.0 (IAX2) suffers from two security holes. Firstly, attackers can carry out Denial of Service (DoS) attacks against Asterisk servers. Secondly, the attacker is able to spy on accounts for which no or only weak passwords exist. After gaining access to such accounts the Asterisk server can be used for DoS attacks through UDP flooding. This compromises the Internet connection of the victim making services unavailable. Although patches do exist for the above mentioned security holes IAX can not be regarded secure for the time being.

3.2 Potential threats and attacks

As attack tools are easily handled even by non-experts and due to their rapid distribution via the web, they constitute a growing danger.

Besides standard attacks against networks and networked IT-systems there are specific attacks against VoIP systems. These threats pertain to all **network layers**.

VoIP service availability depends directly on network infrastructure availability. Therefore, attacks such as Denial-of-Service (DoS) can tear down VoIP links within seconds.

Since VoIP uses TCP and UDP it is sensitive to low-level-attacks such as

- Denial-of-Service (DoS)
- ARP, MAC, IP, UDP, IRDP spoofing
- SYN-, PING- oder MAC- Flooding
- TCP-Session-Hijacking
- RST-Attack
- Data Injection through ISN-Guessing
- Sniffing
- Replay

These attacks become even easier if network zones share the same trust level without user authentication.

On the other hand there are the following attacks against the **application layer** take into account:

1. **Toll interception:** In contrast to eavesdropping PSTN calls, which require physical access to the communication between partners, malware such as trojans are sufficient to sniff and copy speech

packets and to even send them to someone else. Voice mailboxes can also be compromised this way.

2. **Manipulation of calls:** By means of a MitM attack speech packets of a call can be selectively modified.
3. **Unauthorised usage/phreaking/toll fraud:** If an attacker is able to compromise user credentials (VoIP provider access credentials) he can set up calls at the expense of the user (toll fraud). He could even pretend to be a valid user, which is denoted by "phreaking".
4. **Dialer:** Softphones are exposed to a particular risk, since trojans or worms are able to autonomously establish calls without any user notice. This problem can be solved by blocking out numbers or by arranging for value added services by the respective VoIP provider.
5. **Violation of Privacy:** Credentials and other user (subscriber) information can be collected with the aim to monitor and analyse communication profiles.
6. **SPIT (Spam over IP Telephony):** Comparable to Spam-Mails, SPIT massively sends VoIP messages, usually for advertisement purposes. However, SPIT is not much prevalent yet, but will be increase in the future.

Further **security risks** can be named as dynamic port usage, configuration of network devices etc.:

1. **Dynamic port usage:** A multimedia call over IP typically consists of two paths: A signalling path and a media stream path. Typical IP telephony systems use different protocols for call signalling (SIP or H.323) and media communication using RTP (Real-time Transport Protocol). RTP packets are routed by destination IP address and port number for each call and the IP telephony system needs to use a wide range of port numbers. Four ports within this range are required per connection, two for each path. RTP and RTCP dynamically assign ports within the range of 1024 to 65535 and the Session Description Protocol (SDP) transmits port information in the signalling messages. In addition, ports vary between each connection (session). Both paths have problems in traversing networks with firewalls and/or Network Address Translation (NAT). Opening this range endangers any network.
2. **Configuration of network devices:**
   - **Default Ports:** Some users do not modify the default configuration of network devices having many open ports (default ports), therefore, making them vulnerable for DoS attacks, buffer overflow exploits or similar attacks.
   - **Passwords:** Default or trivial passwords for network device configuration/administration are also common risks. Crackers can easily seek out passwords using default password lists and dictionary attacks.
   - **Administration:** It is not recommended to use HTTP and Telnet for administration, since these protocols transmit credentials in plaintext, i.e. unencrypted.
3. **Faulty implementation of VoIP protocols:** Implementation flaws (programming mistakes) of protocols rapidly attract hackers to find security holes or anomalies (crashes, resource leaks) in devices. For example inadequate checking of the size of a protocol request results in several exploits such as DoS attacks or remote access.
4. **Attacks against IP PBX:** As IP PBX are the primary components in VoIP, providing services on the IP network they are very likely to be attacked. Operating systems and the software itself are targeted.
5. **Attacks against operating systems in VoIP systems:** VoIP security also depends on operating systems vulnerability. To give an example: Cisco's Call Manager runs on Microsoft Windows and the Avaya Tenovis Communication Manager runs on Linux. Exploits have been identified for both operating systems, enabling full administrative system access by using buffer overflow attacks.

Furthermore, there are many **attack tools** available to attack VoIP systems directly. In the following some of the most common attack tools and also those operating at network level are shortly listed. A comprehensive overview is given in [9], which provides categories, descriptions and links to current open source, free security tools. VoIP users can inform themselves about attack tools and gain a broader view on how to defend their VoIP devices and deployments.

The following attack tools address vulnerabilities of SIP and RTP. For example, RTP is vulnerable to an attack that congests the network or slow down network end-points, i.e. devices. Attackers having access to the network segment where the media packets are communicated can place a large number of RTP packets into the network causing excessive overload due to massive traffic.

- **Cain & Abel:** Cain & Abel uses ARP spoofing and ARP poisoning. This tool enables sniffing and recording of VoIP conversations. It supports a big amount of different password formats.
- **Vomit:** Vomit converts a Cisco IP conversation into a Wave-File. It requires a tcpdump output file Vomit only supports G.711 coding.
- **VoIPong:** VoIPong finds all conversations in a network, which are coded in G.711. It supports SIP, H.323, Cisco skinny protocol, RTP and RTCP. Like Vomit, it converts conversation to a wave-file.
- **SiVuS:** SiVuS is a vulnerability scanner for VoIP networks. It comprises three components:
  - The SIP Message Generator to test issues or generate demonstration attacks.

  - The SIP Component Discovery to identify targets for analysis.
  - The SIP Vulnerability Scanner to verify the robustness and security of SIP phones, proxy servers and registrar servers.
- **SIPcrack:** SIPcrack is a SIP protocol login cracker. It consists of two programs:
  - SIPdump to identify logged in SIP users.
  - SIPcrack to crack the passwords of the SIP users by means of bruteforce attacks.
- **RingAll:** This simple mechanism allows for a DoS-attack against unsecured SIP clients. It sets the field "User-Agents" with the value "RingAll" and thus forces a broadcast call.

The following attack tools operates at networks level:

- **Wireshark:** This is a network packet analyser which monitors and records transmitted packets for analysing purposes. It can be misused for attacks such as sniffing of user credentials during connection establishment. It supports SIP and H.323 and records conversations in the ".au" file format.
- **Sipsak:** Sipsak is a small command line program for SIP developers and SIP administrators. It can be used for simple tests of SIP applications.
- **Nmap:** This port scanner is normally used for port validation of hosts. It is possible to identify active hosts and open ports in an operating system. With this tool an attacker can easily make use of weaknesses within an operating system.
- **THC-Hydra:** THC-Hydra is a logon cracker which supports numerous protocols such as SIP. It is a proof-of-concept-tool which cracks the password of a specific protocol.

3.3 Risk assessment and implications on security

The following table summarises the implications on the overall security of some of the above described threats. The first table shows general threats and their implication of the main security requirements integrity, authenticity, confidentiality, and availability. On the other hand, the second table shows named attacks directly. Here, the most attacks want to destroy the availability of VoIP systems and are not spying tools regarding recording or evaluation of voice and data streams. But there are many more attacks possible by using standard networks and protocols as in PSTN networks before.

Additionally, the majority of consumer VoIP solutions do not support encryption yet. As a result, it is relatively easy to eavesdrop on VoIP calls and even change their content. There are several open source solutions that facilitate sniffing of VoIP conversations. A modicum of security is afforded due to patented audio codecs that are not easily available for open source applications, however such security through obscurity has not proven effective in the long run in other fields. Some vendors also use compression to make eavesdropping more difficult. However, real security requires encryption and cryptographic authentication which are not widely available at a consumer level. [11]

Attacks	Integrity	Authenticity	Confidentiality	Availability
Disturbing the normal course of operations	√			
Subscriber unreachable	√			
Eavesdropping conversation data*		√	√	
Sniffing registration data on VoIP servers or gateways*		√	√	
Manipulating modifying data*		√	√	√
Hijacking connections or sessions*		√		√
Identity fraud*		√		√
Circumventing communications*	√			
Toll fraud*				√
Interfering the QoS*				√
Malfunction of devices*	√			

Table 1: Threats and their implication on the overall security (part 1), * = Redirecting data streams

Attacks	Integrity	Authenticity	Confidentiality	Availability
DoS /DDos				√
MAC, Ping, SYN, LAND Flooding				√
TLS Connection Reset				√
Replay Attack	√	√	√	√
DHCP Starvation Attack				√
MAC-Spoofing		√	√	
ARP-Spoofing	√	√	√	√
IP-Spoofing		√	√	√
DNS-Spoofing	√	√	√	√
Password Sniffing	√	√	√	√
SPIT				√

Table 2: Threats and their implication on the overall security (part 2)

4. Security mechanisms and standards

In the following, we discuss security mechanisms and standards for SIP and H.323 based infrastructures and applications alike:

4.1 SIP Digest

This digest authentication algorithm (RFC-2617) is currently the most frequently deployed security mechanism

with SIP. Derived from HTTP Digest it allows authentication of a SIP subscriber (user agent, proxy-server or registrar server). It is based on the transmission of a shared secret, which consists of a checksum over a nonce and parameters (user name, password, nonce, SIP method, Request URI). SIP Digest does not exchange passwords. The shared secret is hashed using MD5 or SHA-1. SHA-1 is the IETF recommendation.

4.2 SIPS

SIPS (SIP over SSL/TLS[1]) protects sensitive data such as SIP URI, IP addresses from sniffing or message manipulation. The URI scheme slightly differs from the conventional SIP URI: *sips:this_is_me@sip.com.* SIPS encrypts a connection between a SIP subscriber and a SIPS URI and the network instances (user agent, proxy server, DNS server, location server) in between via SSL/TLS. However, because of SSL/TLS SIPS has to be transmitted via TCP, instead of UDP. The default TCP port for SIP over TLS is 5061. User authentication is accomplished by SIP Digest, which hashes the SIP message (digital signature).

4.3 SRTP

As RTP and RTCP do not offer any protection against sniffing and manipulation of VoIP data, SRTP (Secure Real-Time Transport Protocol; RFC-3711) has been developed. It constitutes an alternative to IPsec-based VPN communications, particularly for real-time transport. SRTP encrypts data symmetrically with AES[2] (Advanced Encryption Standard), i.e. SRTP is the secured variant of RTP and SRTCP of RTCP respectively. They complicate attacks such as sniffing, replay and DoS. For transportation RTP/RTCP packets are encapsulated in SRTP/SRTCP packets. Security features of SRTP comprise: [2]

- Encryption of the media stream (against sniffing)
- Authentication of the sender (against identity spoofing)
- Validation of the integrity (against modification/manipulation)
- Replay[3] protection (against unauthorized access to end-points)

SRTP defines two kinds of keys: Master Key K_M and Session Key (K_E for encryption and K_A for authentication). K_E (minimum 128 bit) und K_A (minimum 160 bit) are derived from the K_M by means of a cryptographic secure pseudo-random function (PRF).

Encryption is applied to the RTP data stream and key exchange is accomplished via signalling, which means, that a secure key exchange needs secured signalling, e.g. SIPS.

There is no inherent mechanism for master key generation and management. Multimedia Internet Keying (MIKEY; RFC-3830) is viable to be applied here. MIKEY describes key management for real-time multimedia communications and allows key exchange (Transport Encryption Key (TEK) and Transport Generation Key (TGK)) and further security parameters (Data Security Association) between subscribers. MIKEY serves for the exchange of the Master Key K_M and security parameters. As the TEK can be updated (re-keying), K_M can be updated as well. MIKEY provides peer-to-peer and one-to-many communications and is dependent on the underlying protocol, i.e. SIP and H.323. This means, that one end-point with SIP can establish a secured communication link with an end-point that uses H.323. Furthermore, as MIKEY is able to use different process key exchanges and security parameters for different sessions in parallel, RTP and RTCP links can be secured independently and concurrently even with a shared TEK for all. [3]

4.4 S/MIME

S/MIME (Security/Multipurpose Internet Mail Extension; RFC-2311) has been originally designed to encrypt and authenticate message bodies (MIME) in email communications. However, MIME is not restricted to email messaging. It can be applied for secured end-to-end transport of message bodies within IP and thus it is suitable for SIP. In addition to SDP parameters detailed subscriber information data (e.g. presence information) can be secured as well.

In contrast to SIP, where encryption is only applied on a „hop-by-hop"-basis and thus information contained in the message bodies are unencrypted within the traversed SIP network, S/MIME offers „end-to-end"-encryption. Information in the message bodies can exclusively be read by the end-points by means of asymmetric cryptography. Key exchange (of public keys) can either be done by SIP or through certificates.

4.5 H.235

From H.323 (ITU-T) version 4 on the structure of the substandard H.235 for H.323 security changed from annexes (D to I) to documents from H.235.0 (framework, a.k.a. H.235v4) to H.235.9. Now in H.323 version 6 security features have been extended. Support for SRTP is the most essential improvement [10]. The following security features are implemented:

- Subscription based authentication (through symmetric cryptography)
- Authentication by means of certified public keys
- Diffie-Hellman (DH) key exchange (DH key negotiation with public keys, generation of a symmetric key for authentication)

For the time being the only vulnerability known in H.323 is the possibility to take advantage of ASN.1 parsing defects in the first phase of H.225 (initial call setup) data exchange. Furthermore, complexities of the protocol suite and the high number of different vendor implementations

[1] SSL/TLS Usage within SIP
[2] AES-CTR and AES-f8
[3] In a replay attack an attacker resends recorded RTP or RTCP packets. This type of attack can cause DoS attacks. Only integrity protection through message authentication can protect from replay attacks. SRTP has this ability.

(ASN.1/PER[4] Coding/Decoding) [5] constitutes an additional problem.

Through Media Anti-Spam it is possible for the receiver to check, if a RTP packet is authentic and is originated by an authorised sender.

4.6 ZRTP

ZRTP, as implemented by ZFone, has been developed by Phil Zimmermann in order to achieve interoperability between SIP end-points from different vendors. ZRTP is designed to provide authentication between parties, secrecy, and perfect forward secrecy between sessions. ZRTP is an extension of RTP and describes a key agreement/establishment protocol for use with SIP and SRTP without the need for a shared secret or a separate public key infrastructure (PKI) – instead voice authentication digests can be used to verify identities.

It uses Diffie-Hellman key exchange during call setup "in-band" in the RTP stream. ZRTP packets are embedded in RTP packets. The initial DH exchange generates the shared secret from which the Master Key for SRTP sessions is derived, which is passed on to the SRTP layer. The SRTP layer derives Session Keys from the Master Key material and handles replay attack protection via authenticated sequence numbers. AES is used for payload encryption and the SRTP packet, including headers, is authenticated using SHA-1. The protocol assumes that the call has been established using a signalling mechanism such as SIP. Each SRTP instance is identified by its unique ZID (ZRTP Identification Data), a 96 Bit random value which is being created for an installation of Zfone. If previously shared secrets exist, ZID allows for ascertaining.

4.8 SPIT filtering

A countermeasure to SPIT attacks is SPIT filtering with different mechanisms like buddylists/whitelists and blacklists. I.e., each VoIP subscriber has a list containing the subscribers he wants to communicate with or not. The whitelist mechanism is an efficient approach but not very practical, since subscribers can not be reached if they are not listed. However, there are also extended whitelists with a web of trust.

4.9 Assessment

SIP Digest features several major weaknesses which can be easily exploited. One of the weaknesses is the limited message integrity (the header is not included in the integrity calculation) [7]. An attacker can easily change the message or can be a MitM sniffing valid credentials, change them and send it to a server. Since most implementations accept the same credentials within a period of time, the attacker could replay messages (replay attack). In this context an attacker can register as a legitimate user. He is also able to redirect conversations to his device. Furthermore, some request methods do not use the digest algorithm.

Due to the transactional model in SIP the request methods CANCEL and ACK are weakly authenticated. This is more or less impossible, since these methods operate in hop-by-hop mode and thus may be generated by any instance (server) in the signalling chain. It is improbable that every server has a security association with other instances, making authentication of these requests is effectively impossible. Also, the sequence numbers of these two request methods must be the same as the one of the requests to which they relate and thus cannot be challenged (leading to incrementing the number and thus not matching with the original message).

This lack of authentication of CANCEL and ACK enables attackers to carry out injection attacks. An attacker can fake a CANCEL request resulting in a denial of session establishment. He can create a malicious ACK message (credentials in an ACK message are identical with those of the previous request) [7]. Dorgham Sisalem et. al [7] propose an approach to mitigate the security risk by using predictive nonces. This technique allows cryptographical binding of critical header fields to challenges, thus preventing attackers from changing them and guaranteeing their integrity.

SIP has a number of security mechanisms, mentioned in **SIPS**. Some of them have been built in to the SIP protocol directly, such as HTTP authentication. These mechanisms have alternative algorithms and parameters. SIPS does not provide end-to-end-security, but secure hop-by-hop-communications. This requires network instances to authenticate each other. RFC-3261 itself does not provide any mechanism agreement options. Moreover, even if some mechanisms such as OPTIONS were used to perform a mechanism agreement, the agreement would be vulnerable to Bidding-Down attacks[5]. Three header fields are defined for negotiating the security mechanisms within SIP between a SIP User Agent entity and its next hop SIP server. It is a proposed standard (RFC-3329) from the IETF. Five mechanisms are currently supported:

- TLS
- HTTP Digest
- IPsec with IKE
- manually keyed IPsec without IKE
- S/MIME

Currently there are two drafts being discussed within the IETF dealing with end-to-middle, middle-to-middle and middle-to-end security: "End-to-middle Security in the Session Initiation Protocol (SIP)" [13] and "A Mechanism to Secure SIP information inserted by Intermediaries" [14]. The security requirements between both approaches are slightly different, since information is added by intermediaries and used by intermediaries. Nevertheless, SIP End to Middle Security [13] and SIP Intermediate Security [14] share the same fundamental problems to be solved in SIP. [12]

[4] PER: Packet Encoding Rules

[5] a phase of man-in-the-middle attack where the attacker modifies messages to convince communicating parties that both sides support only weak algorithms

Security mechanism	Confidentiality	Integrity	Authentication	Access Control	Liability	Anonymity
SIP Digest	-	-	+	+	+	-
SIPS	+	+	(+)	-	-	-
S/MIME (Message Body)	(+)	(+)	(+)	(+)	(+)	(+)
SRTP	(+)	(+)	(+)	-	-	-
H.235	+	+	+	-	-	-
ZRTP	(+)	(+)	(+)	-	-	-
IAX2	+	+	+	-	-	(+)
Skype	(+)	(+)	-	-	-	-

Table 3: Overview about VoIP security protocols

As an end-to-end protocol **SRTP** does not dependent on the network infrastructure and thus is suitable for public networks. Cryptoanalysis is reduced by re-keying of master and session keys. Due to its low resource consumption it has no qualitative impact on a VoIP service.

SRTP provides mechanisms against replay attacks for the receiver to manage "replay lists" which contain indexes of earlier received authentic packets. The receiver can analyse a newly received packet regarding collisions. Therefore, the memory of IP phones should be of a appropriate size.

Authentication and integrity of RTP messages are realised by HMAC-SHA-1 with K_A. Because of the weaknesses of HMAC-SHA-1, it is recommended to use SHA-256, though not yet standardised. Some experts recommend encrypt RTP payload and then calculate the fingerprint of the encrypted payload.

SRTP enables secure RTP sessions during connection establishment phase, which is accomplished prior to SRTP packet exchange. However, sensible data (e.g. Call-ID, From, To, Via, Codecs) are communicated in plain-text. Attacks such as MitM, spoofing or phishing are possible. This requires protection of the data during the connection establishment, i.e. encryption of SIP packets.

SRTP may be also vulnerable to Bid-Down (MitM) attacks. The attacker forces a lower encryption level (e.g. from AES-256 to AES-128) by removing the information AES-256 in the INVITE message. [6]

One problem arises with **S/MIME**: There is no organisation that manages worldwide distribution of certificates; i.e., there is no global public key infrastructure (PKI). Furthermore, S/MIME secured connections take to long to establish a session.

For the time being the only vulnerability known in H.323 is to gain advantage of ASN.1 parsing defects in the first phase of H.225 data exchange. **H.235** will improve that. Furthermore, complexities of the protocol suite and the high number of different vendor implementations (ASN.1/PER[6] Coding/Decoding) [5] constitute an additional problem.

Though key exchange is not protected against MitM-attacks, the communication is secured. Given normal conditions **ZRTP** is cryptographically secure. However, this protocol is vulnerable to adversaries with strong capabilities. Some experts advise for modifying the protocol to include a randomised start-time for the conversation.

IAX is the Inter-Asterisk eXchange protocol used by Asterisk, an open source PBX server from Digium. It is used to enable VoIP connections between Asterisk servers, and between servers and clients that also use the IAX protocol. IAX now most commonly refers to IAX2, the second version of the IAX protocol. The original IAX protocol has been deprecated almost universally in favor of IAX2.

IAX2 is a very robust and full-featured yet simple protocol. It is agnostic to codecs and number of streams, meaning that it can be used as a transport for virtually any type of data. IAX2 uses a single UDP data stream. usually on port 4569, to communicate between endpoints, both for signalling and data. The voice traffic is transmitted in-band, making IAX2 easier to firewall and more likely to work behind network address translation. This is in contrast to SIP, which uses an out-of-band RTP stream to deliver information.

The basic structure of IAX is that it multiplexes signalling and multiple media streams over a single UDP stream between two computers. IAX is a binary protocol, designed to reduce overhead especially in regards to voice streams. Bandwidth efficiency in some places is sacrificed in exchange for bandwidth efficiency for individual voice calls. One UDP stream is easier to setup for users that are behind a firewall. Additionally, IAX2 can use the encryption mechanisms AES-128 to hide signalling and voice data on a secure way.

Skype users essentially make telephone calls and video calls through their computer using Skype software and the Internet. The basis of the system is free communication between users of Skype software; however the product also allows Skype users to communicate with users of regular landline and mobile telephones. This software is currently available free of charge and can be downloaded from the company website, but the software is proprietary and the Skype protocol is unpublished.

The main difference between Skype and other VoIP clients is that Skype operates on a peer-to-peer model, rather than the more traditional server-client model. The Skype user directory is entirely decentralised and distributed among the nodes in the network, which means the network can scale very easily to large sizes (currently just over 100 million users) without a complex and costly centralised infrastructure.

Skype generates a significant amount of discussion on how secure its traffic really is. It has had an impact upon the

[6] PER: Packet Encoding Rules

security and culture of VoIP telephony because of this discussion and several design principles:

- All Skype traffic is encrypted by default and the user cannot turn it off
- Skype reportedly uses openly available, strong encryption algorithms
- The user is not involved in the encryption process and therefore does not have to deal with the issues of public key infrastructure.

This has had an effect upon the rest of the market as they seek to offer competitive products. The security of Internet communication has become an issue of which people are more aware and secure communication a feature they want to see in the products they use.

5 Security measurements and concepts

Since VoIP networks consist of a wide spectrum of different components like phones, call managers, gateways, servers, routers, etc. VoIP infrastructures demand specific requirements regarding security. Overall security has to comprise both standard network security and VoIP specific security parts.

In the following some basic security measurements, common concepts and specific recommendations for securing VoIP applications and VoIP-based infrastructures in enterprise environments are discussed.

5.1 Standard network security

Protection of a VoIP network has to start with standard network security measures applied for data networking. Additionally, new techniques are necessary due to special and emerging risks associated with VoIP. A prerequisite and essential preparation is an in-depth planning to insure reliable service with a satisfactory quality-of-service (QoS).

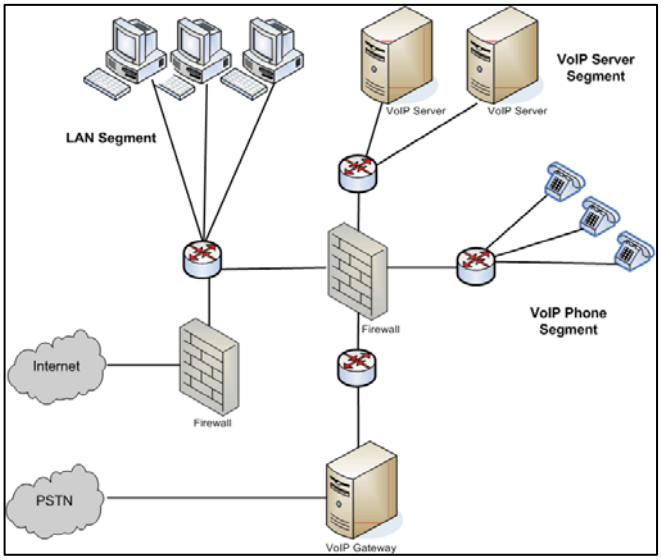

Figure 2: VLAN and firewalling

It is important to bear in mind that VoIP must not be perceived as another IP application. While data network security primarily focuses on the data network layer by means of e.g. VLANs and firewalls VoIP demands a multi-layered approach, comprising the identification of physical devices (phones, gateways and servers) and networks entry points, together with their authorisation. In the next step, the "session layer" and the "transport layer" have to be addressed, since they may be exposed to typical attacks such as DoS and MitM. Finally, layer 7 in the ISO-OSI model – the application layer – has to be protected from viruses, SPIT attacks, toll fraud, identity theft, etc.

5.2 Virtual LAN (VLAN)

It is highly recommended to establish two separate VLANs – one for data and one for voice traffic – as ideally separated subnets with different RFC-1918 address range (10.x.x.x, 172.16.x.x, and 192.168.x.x). In addition, deploying separate DHCP servers would ease intrusion detection and firewall protection. The logical separation ensures that both data network and VoIP network can not be compromised. Viruses can not infect both sides of the network. Also it is much more difficult for an attacker to sniff, intercept, or eavesdrop on traffic.

A further advantage of a VLAN is the possibility for providing priority classes for quality-of-service management. Speech data can be routed with a higher priority than normal data.

In the VoIP VLAN only phones may operate and therefore network administrators have to monitor this network to avoid potential compromising by rogue phones.

In addition, through separation of voice and data traffic into separate collision domains there is no data competition thus reducing latency (queue/wait time) for transmission services. Audio streams are very latency sensitive, so this low-cost approach improves performance in an existing network infrastructure.

However, even if VLANs constitute an elegant way, a complete separation without VLAN routing could be problematically, because of the following reasons: [15]

- Softphones at workstations within the data network need access to the VoIP server in the VoIP network
- Groupware clients need direct dialing of contacts from an application (e.g. Lotus Notes or Outlook)
- The VoIP server is connected to directory services such as LDAP or ADS

Also, some enterprise infrastructures are too complex for VLANs because of the number of different systems such as VoIP server, VoIP gateway, VoIP phones, softphones, unified messaging server, CTI systems, etc.) connected with each other. These problems can be solved if the VoIP server is standing in the voice and data VLAN simultaneously.

Figure 2 depicts a viable topology with two firewalls – one for the communication to/from the Internet and one firewall via PSTN. The first one only routes packets from the Internet to the internal network, if they belong to a session initiated from the internal network. The second firewall separates the network segments LAN, phone, VoIP server, and VoIP gateway and controls the traffic between these segments.

5.3 Encryption

Encryption allows privacy and authentication in phone communications. Securing call streams is possible through SRTP, encrypting RTP content, and call signalling with

TLS. TLS is an alternative to IPsec and provides effective security against hijacking attacks whereas SRTP preventing eavesdropping attacks.

IPsec is used to encrypt call setup and control messages. It is an effective measure against eavesdropping and protects sensitive information. However, devices and call managers have to be checked respectively, if they support these protocols[7].

Besides, there may be interoperation problems with some management and monitoring tools, since they do not deal with encryption.

IPsec should be used in "tunnelling mode" instead of "transport mode", since tunnelling masks the IP addresses (source and destination) and thus secures the communication path against traffic analysis. Also there is a problem with NAT. IPsec uses port 500 which circumvents multiple VoIP tunnels through a NAT.

The majority of consumer VoIP solutions do not support encryption yet. As a result, it is relatively easy to eavesdrop on VoIP calls and even change their content. There are several open source solutions that facilitate sniffing VoIP conversations. A modicum of security is afforded due to patented audio codecs that are not easily available for open source applications, however such security through obscurity has not proven effective in the long run in other fields. Some vendors also use compression to make eavesdropping more difficult. Real security requires encryption and cryptographic authentication which are not widely available at a consumer level.

5.4 Authentication

Unlike PSTN systems, where each phone is given its own phone number (ID) which is matched to the physical location of the phone at the line, VoIP systems assign phones an IP address. As these can be spoofed by an attacker, peer-to-peer authentication is mandatory. However, this is difficult to realise as secret information needed for such authentication can not easily be shared between the users making the need for a PKI-based approach.

Privacy between users communicating (talking) on a VoIP network is of major concern, since the shared network on IP basis can be accessed by an attacker eavesdropping conversations by means of a MitM-attack. Appropriate encryption and authentication of conversation are required.

Since a compromise of central VoIP systems such as VoIP gateways or VoIP server would jeopardise the entire VoIP network and its components and thus the entire communication of the enterprise they have to be protected from unauthorised access. Hence, users and administrators need to be authenticated by a central authentication service, which is connected to a firewall system. The firewall system itself has to control the limited access.

5.5 Firewalls

Firewalls as standard security feature on data networks inspect each packet that travels to and from the network. SIP and H.323 based IP telephony services use UDP packets and incoming TCP connections. However, most firewalls in enterprise environments are configured in such a way that they do not allow them to pass through. Furthermore, due to dynamic port assignments throughout the call, firewalls have problems with filtering VoIP traffic. It does not matter which signalling protocol is used, SIP or H.323. Both protocols necessitate stateful inspection filtering to track traffic and associate port numbers. This kind of firewall is able to remember previous traffic and can investigate the application data in a packet. It can track the state of a connection and deny packets from an unwanted call origin.

Application level gateways (ALGs[8]) are an alternative. Since they understand the VoIP protocol data carried as payload in an ordinary packet they are "VoIP-aware". ALGs contain software which are parsers for ASN.1 (H.323 is coded in ASN.1), SIP, MGCP and SDP temporarily record the states of the signalling protocols and dynamically open/close ports according to the session state.

Compared to stateless and stateful firewalls, ALGs offer the highest security level, since they only open UDP ports actually needed for the duration of a communication instead of opening a range of ports.

Even if some ALGs for VoIP protocols are available on the market, most of the existing firewalls cannot efficiently handle VoIP protocols such as SIP. Also vendor proprietary protocols relay on dynamic port ranges and do not support NAT. Session Border Controllers (SBC), a new generation of firewalls, are addressing most of these problems. SBCs control both signalling and media traffic (speech data, video, data).

The firewall should be able to monitor signalling and VoIP-aware NAT and media session management.

When implementing Campus VoIP, the VoIP network or segment should be prevented from transporting data to and from the Internet by means of a firewall. In addition, the number of calls has to be limited over the WAN traversing through the media gateway in attempt to prevent DoS attacks.

5.6 Intrusion Detection Systems and Intrusion Prevention

IDS/IPS are quickly becoming an integral part of most firewall-based security architectures. They block malicious packets and protect the network system against intrusion. However, misconfigured or not well-tuned inline IPS may drop normal packets, since they erroneously detect these as malicious packets (false positive). The typical small size of UDP packets can be interpreted as a UDP flooding DoS

[7] SRTP is available on Analog Telephone Adapters (ATAs) from some vendors like Sipura/Linksys, with a certificate from Voxilla, a VoIP information site. SRTP is also available for Gizmo Project for softphones (PCs/laptops emulating a phone).

[8] An ALG is an application-aware entity that examines a particular application protocol flow and only allows messages that conform to a security policy. It may also modify messages so that they will conform to the policies and be able to pass through.

attack. This can result in re-transmission causing loss-of-service disconnects. Therefore, VoIP-aware and VoIP-enabled border IPS should be used.

5.7 NAT and STUN

NAT (Network Address Translation; RFC 3022) is a typical feature on a network which translates private IP addresses into public ones. NAT routers only pass packets if the connection has been initiated from in the inside and spontaneously incoming calls from outside are being rejected. A viable solution is to forward the port (incoming UDP speech data) to the recipient. However, these are associated dynamically and this can be handled in a different way.

Even outgoing calls can impose a problem, since NAT can only translate layer 3-packets, but TCP operates in layer 4. NAT modifies the source IP address (in the IP header) and source port (in the UDP or TCP header), whereas IP address and UDP port information stay unchanged within the signalling message part[9]. This requires keeping the internal IP address as sender address in the IP header of outgoing packets. The result: these calls can not be answered, since the internal IP addresses can not be routed. The public (external) IP address needs to be transmitted to the recipient. STUN (Serial Tunnelling; RFC 3489) helps to solve this problem. With this protocol end-points can get their public IP address and the NAT binding of the gateway. For this, the STUN client (e.g. the VoIP phone) sends a request to the STUN server which then assigns the client's credential information (user name and password). The client sends a second request using the credentials with the aim to receive the NAT binding information of the NAT gateway located in front of the STUN server. The latter extracts the source IP address and source IP port out of the message und sends these within a reply message to the STUN client. Afterwards, the respective VoIP application exchanges public IP addresses with internal IP addresses and inserts these into the header. Further requests allow the client to identify the NAT type. STUN is being supported by numerous VoIP phones and providers.

If possible, NAT should be avoided or mechanisms such as STUN or TURN should be applied.

5.8 Softphones vs. IP hardware phones

Phones are the most common component in a VoIP network. Unfortunately, they are also most targeted by attackers.

Softphones should be avoided in any case. They violate the separation of voice and data and they are vulnerable to malware such as worms and viruses due to the number of possible entries into the system. These entry mediums include the operating system, resident applications, and enabled services. Unlike IP hardware phones, which are placed in the VoIP segment of the VLAN, softphones reside on the data segment and are susceptible to any conventional attack against that entire segment. IP hardware phones run on proprietary operating systems with limited network services.

The majority of VoIP handsets use proprietary, closed protocols such as Unistim (Nortel), SCCP (Cisco) or H.323 with proprietary extensions (Avaya). Proprietary protocols can be more secure, since they are not widely documented and attackers are not able to analyse and exploit vulnerabilities. Also, simple VoIP handsets are more difficult to attack, because they are running less complex software.

It is recommended to use phones offering strong security (authentication and/or encryption) for signalling and media. Furthermore, they should obey security policies. Remote access features such as TELNET may not be allowed.

5.9 Network devices

It is highly recommended to use cryptographically strong passwords. If HTTP or TELNET is to be used for administration, the connection between the client and the network devices should be secured with SSL/TLS or SSH.

5.10 Operating systems

VoIP systems use general-purpose operating systems, which tend to have more vulnerabilities than proprietary ones. A principal protection mechanism is to shut down all unused services in an operating system of a VoIP system, such as VoIP server or IP PBX.

Additionally VoIP systems running on common operating systems such as Windows or Linux have to be hardened. Furthermore, unnecessary services have to be disabled.

5.11 QoS (Quality-of-service)

It is necessary to take into account that firewalls such as stateful firewalls or ALGs can impose variable latency on traffic, with strong impact on QoS. Also, other VoIP systems such as IP phones have low computational power to perform encryption and only few IP phones provide AES encryption at reasonable cost. In this case, it is recommended to place encryption at a central point such as router or gateway, which is responsible for the encryption of all traffic from any host or system, e.g. by IPsec tunelling.

Also some protocols impose implications on QoS. SIP for example encodes messages in ASCII format, which are sometimes too large for low-bandwidth networks. As they can sometimes exceed the MTU size over Wireless LAN links, they can cause delay or packet loss. An approach to alleviate this problem is to binary encode SIP as proposed by RFC-3485 and RFC-3486.

Availability is a further issue. Additional power backup systems can ensure continuity in operation during power outages.

5.12 Remote management

It is recommended to use IPsec or Secure Socket Shell (SSH) for all remote management and auditing access purposes. A further possibility is to use HTTPS with TLS. Remote management should be avoided for access to IP-

[9] Session Description Protocol inserts the port information into the signalling message part.

PBX systems should be accomplished from a physically secure system. If remote management is necessary it should be use from the same Intranet and every time with encryption. Additionally it can be used for remote access from other participants from the same company.

5.14 Patches

Network administrators have to be especially diligent about patching current versions of VoIP software or firmware.

6. Best-practice approaches for minimising common VoIP network risks

The main open issues of VoIP are quality-of-service (QoS) and security. Regarding security, this paper shows how possibilities are available and which risk can be appearing. But, SIP and RTP are improved regarding encryption and authentication and also H.323 can provide a secure communication. It is a question of implementation into the VoIP equipment and the knowledge about it. Additionally, there are further developments on work which will improve the VoIP technology in the next future. The following table summarises best-practice approaches for typical VoIP risks at the end of this paper to give a last overview.

Risk	Best-Practice Approach
Application-level attacks	• ALGs, firewalls and application-aware IDS/IPS
DoS/DDoS	• Application-aware IDS/IPS • Maintain current patch levels • Antivirus system • Policy-based security zones • VLAN
Eavesdropping	• VPN to isolate VoIP traffic • Selective encryption
Attacks against protocols	• ALGs and IDS/IPS
SPIT	• Strong authentication, authorisation and IPsec
Unauthorised SIP monitoring, spoofing	• Strong authentication, authorisation and IPsec
Viruses and worms	• Current patch levels • Antivirus system • Application-aware IDS/IPS • Policy-based security zones • VLAN

Table 4: Best-practice approaches for typical VoIP risks

7. References

1. www.packetizer.com/voip/h323/whatsnew_v6.html
2. BSI – VoIPSEC: Studie zur Sicherheit von Voice over Internet Protocol, page 35, www.bsi.bund.de/literat/studien/VoIP/index.htm
3. BSI – VoIPSEC: Studie zur Sicherheit von Voice over Internet Protocol, pages 109-111, www.bsi.bund.de/literat/studien/VoIP/index.htm
4. Evren Eren, Kai-Oliver Detken: Mobile Security - Risiken mobiler Kommunikation und Lösungen zur mobilen Sicherheit. Carl Hanser Verlag. ISBN 3-446-40458-9; München Wien 2006
5. www.securityfocus.com/infocus/1782
6. www3.ietf.org/proceedings/06mar/slides/raiarea 1/raiarea-1.ppt
7. Dorgham Sisalem et. al., SNOCER, Low Cost Tools for and High Available VoIP Communication Services, Towards a Secure and Reliable VoIP Infrastructure, 3rd May 2005, pp. 38-39, www.snocer.org/Paper/COOP-005892-SNOCER-D2-1.pdf
8. www.ietf.org/internet-drafts/draft-guy-iax-02.txt
9. voipsa.org/Resources/tools.php
10. www.packetizer.com/voip/h323/whatsnew_v6.html
11. Anderson, Mark: VoIP Security – Uncovered; WhiteDust Security; http://www.whitedust.net; Retrieved on 2006-05-26
12. Kuhn, Walsh, Fries: Security Considerations for Voice Over IP Systems; Recommendations of the National Institute of Standards and Technology (NIST); NIST Special Publication 800-58; January 2005
13. G. Egeland: Introduction to IPsec in IPv6; Eurescom; http://www.eurescom.de/~publicwebdeliverables/P1100series/P1113/D1/pdfs/pir1/41_IPsec_intro.pdf
14. Cisco Networkers 2000: http://www.cisco.com/networkers/nw00/pres/2403.pdf
15. BSI – VoIPSEC: Studie zur Sicherheit von Voice over Internet Protocol, page 86, www.bsi.bund.de/literat/studien/VoIP/index.htm