SIP Security
Status Quo and Future Issues

Jan Seedorf

Security in Distributed Systems (SVS)
University of Hamburg, Dept. of Informatics
Vogt-Kölln-Str. 30, D-22527 Hamburg
seedorf@informatik.uni-hamburg.de

Abstract

Today, the session initiation protocol (SIP) is the predominant protocol for Voice-over-IP (VoIP) signalling. The intention of this paper is to present an overview of VoIP security issues - both current and future – focusing on SIP. We start by presenting some fundamental differences between VoIP and the public switched telephone network (PSTN). We then look at specific problems for SIP signalling that arise from these differences. We summarize current activities regarding SIP security, including recent developments in the research community and standardization efforts within the IETF. Finally, the paper will give an outlook on security issues in future VoIP scenarios. Specifically, we present a short security analysis of using SIP in a peer-to-peer setting (P2P-SIP).

1. Introduction

In recent years, the digitized transmission of audio signals over IP-based networks - commonly called Voice-over-IP (VoIP) - has emerged to a widely used application. During this evolvement the Session Initiation Protocol (SIP) [1] has become a popular and now dominating choice for signalling in VoIP communications. Today, many SIP implementations by various vendors exist (Hardware Telephones, "Softphones", Gateways, etc.). However, due to some fundamental differences compared to the Public Switched Telephone Network (PSTN), VoIP phone calls cannot be considered as secure as phone calls carried out over the PSTN.

The intention of this paper is to present an overview on current and future security challenges in SIP-based VoIP communications[1]. In the next section we describe the differences between VoIP and the PSTN that make securing VoIP difficult. Following this general introduction and motivation on VoIP security we give an overview on signalling with SIP. We then look at current research problems in SIP-based VoIP. We describe some important challenges and give a brief summary on current approaches to solve the problems. Finally, we give an outlook on future SIP-based VoIP scenarios (i.e. Peer-to-Peer SIP) and the security implications of such an - yet another - infrastructure change for VoIP signalling.

2. Differences between Voice-over-IP and the PSTN

VoIP as it is used today has some fundamental differences compared to speech transmission in the Public Switched Telephone Network (PSTN):

- In the PSTN, signalling is done in a separate and closed network. With VoIP, signalling is done in an open, highly insecure network (e.g. the Internet).
- Traditional telephones are simple devices with limited functionality. VoIP terminals, on the other hand, are complex devices with their own TCP/IP stack.
- VoIP offers mobility: users can change their location and still use the same identity in the network. A VoIP-user only needs access to the Internet. By contrast, in the PSTN there is no mobility.
- Because there is no mobility in the PSTN, authentication is not necessary. Anybody who has physical access to a socket in the wall can use that line. As VoIP can be used from anywhere in the Internet, additional authentication must be utilised.

Most security problems that VoIP faces today arise from these significant differences. This is especially true in the consumer market where phone calls are carried out over the Internet. From the perspective of mobility and authentication, VoIP is similar to mobile phone networks such as GSM. However, GSM differs from VoIP because it uses smartcards in terminals and consists of a limited number of providers that trust each other.

[1] We assume that the reader is somewhat familiar with VoIP; a general introduction on Voice-over-IP technology and research challenges can be found in [23].

3. Signalling with SIP

The Session Initiation Protocol (SIP) was specified by the IETF as a standard for signalling and control in multimedia communications over IP [1]. SDP, the Session Description Protocol, is used to select parameters (such as the codec and media type) for the transmission. After a session has been established with SIP, the actual media transfer is transmitted with the Real-time Transport Protocol (RTP). Because SIP is used to set up a session, any secure communication that can be established in a SIP session can further be used to negotiate secrets for a secure RTP stream. Therefore, SIP security is of high importance for VoIP security.

SIP is a client-server protocol which resembles HTTP. Signalling is based on text messages: A message consists of a header and an optional body. Messages are either *requests* or *responses*. If a SIP entity receives a request, it performs the corresponding action and sends back a response to the originator of the request. Responses are three-digit status codes. Table 1 list SIP requests; table 2 lists classes for SIP response codes.

SIP Request	Description
INVITE	*Initiates a call signalling sequence*
BYE	*Terminates a session*
ACK	*Acknowledge*
OPTIONS	*Queries a server about its capabilities*
CANCEL	*Used to cancel a request in progress*
REGISTER	*Used to register location information at a registrar*

Table 1 SIP Requests

SIP Response Codes
1xx – informational
2xx – ok
3xx – redirection
4xx – client error
5xx – server error
6xx – global failure

Table 2 SIP Response Codes

Addressing in SIP is done with Uniform resource Identifiers (URIs). A SIP-URI is similar to an e-mail address and generally of the type "sip:user@domain". SIP designates different (logical) entities: *user agent*, *proxy*, *registrar*, *redirect server*, and *location server*. A *User agent* is a terminal participating in SIP-communications (this can be hardware or software). A *proxy* receives messages and forwards them to another SIP entity. A *redirect server* redirects the sender of the message to another SIP entity instead of forwarding the message. Users can register their current location (i.e. IP-address) with the *registrar* of their domain. This enables mobility: A *location server* is used by a registrar to store the location of users (the binding of a SIP-URI with a current IP-address). The location server provides a directory for other SIP entities to look up the current location for a given SIP-URI.

Example: Setting up a Simple Voice Connection with SIP

The establishment of a voice connection between two users is illustrated in Figure 1 [2]. In this example, user agent A and B are in different domains and have different proxies. First, the callee (user agent B) needs to register with its local registrar (1) to be able to receive calls. The registrar stores the location information at a location server (2). When user agent A wants to call user agent B, it sends an INVITE-request to its local SIP-proxy (3) which passes on the request (possibly after a DNS lookup) to the proxy of user B's domain (4). The proxy in domain B needs lo look up the IP-address of user agent B at the location server (5, 6) before it can send the request to user agent B (7). The response message for user agent A can take the same route back (8, 9, 10).

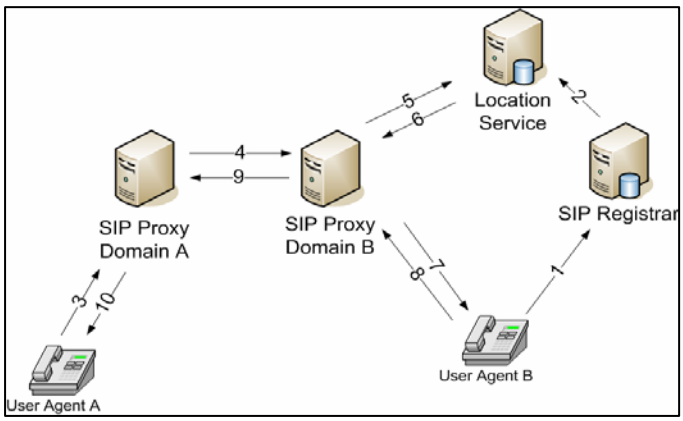

Figure 1 – Setting up a phone call with SIP

SIP Security Mechanisms

The SIP standard, as specified in RFC 3261 [1], includes several security mechanisms:

- S/MIME: Because SIP is using MIME for message bodies, S/MIME can be used to send authenticated and encrypted messages between user agents.
- Digest Authentication: SIP entities sharing a secret (e.g. a password) can authenticate each other with a challenge-response mechanism. To prevent replay attacks, this challenge-response authentication includes nonces.
- TLS & IPSec: Hop-by-hop security for SIP signalling can be achieved either on the transport layer (TLS) or on the network layer (IPsec).

In theory, these security mechanisms can make SIP signalling secure. However, they require a pre-call trust relationship or rely on a trust infrastructure (like a public key infrastructure), which all users can use and with one root that all users trust.

4. Current Security Problems for SIP signalling

By its very definition, VoIP uses IP networks for setting up voice communication. Thus, all threats that are well-known in IP-networks (e.g. denial-of-service, spoofing, sniffing, …) are inherited by VoIP. Furthermore, implementation vulnerabilities (e.g. buffer overflows) are likely because VoIP servers and terminals are complex IP-devices. Specific to SIP are – among others - the following threats (see also [3]):

- Registration/call hijacking
- Impersonating a SIP-entity
- Tampering with message bodies
- Tearing down sessions
- Denial of service
- Eavesdropping
- Spam

Many activities exist with the goal of making VoIP more secure. Within the scope of this paper, it is only possible to list some important challenges and summarize current activities to mitigate these problems. For a more depletive list of threats to VoIP and SIP the reader is referred to [3], [4].

Authentication

One of the fundamental problems for SIP security is end-to-end authentication of communication partners in the absence of a universal trust infrastructure[2]. If the communication partners have a pre-call trust relationship (e.g. via e-mail), S/MIME can be used. Hop-by-Hop solutions (e.g. TLS, IPSec) only work if there is a transitive trust path between sender and receiver of a SIP message. Unlike https, SIP messages via TLS can pass many application layer hops between sender and receiver, and some intermediary entities may not be trustworthy. The following approaches are trying to mitigate authentication problems for SIP/VoIP:

- *ZRTP* [5] is a protocol developed by Phil Zimmermann, the inventor of PGP. ZRTP enables a Diffie-Hellman key exchange within an RTP stream. This key exchange is protected against man-in-the-middle attacks through an authentication string. The user can verify this authentication string with the actual voice of his communication partner. Thus, ZRTP offers authentication of a known communication partner without using any trust infrastructure.
- RFC 3325 specifies a SIP header in which a proxy of a domain can assert the identity used in a SIP message. However, this assertion is not signed. It can be exchanged between domains that have a TLS connection. In [6], a similar "SIP Identity" mechanism is suggested. With this approach, a proxy can assert proper authentication of an identity from its domain and sign such an assertion.
- The SIP community has realised that hop-to-hop security offered by TLS is insufficient for authentication in many cases. The goal of [7] is to develop a new way to establish end-to-end authentication between user agents with SIP.

Security of Terminals & Servers

Because SIP devices are complex, implementation weaknesses seem unavoidable. Vulnerabilities for SIP implementations are found frequently (e.g. [8]). The following efforts strive to make SIP implementations more secure by fostering SIP black-box testing:

- The University of Oulu, Finland, has developed a test-suite for SIP implementations [9]. It contains a large amount of valid and invalid SIP-Invite messages. A test conducted on SIP devices with these messages showed many weaknesses in the tested products [9].
- RFC 4475 describes various test messages that can be used to "torture" a SIP implementation.
- Many simple tools (e.g. [10]) can be used to carry out tests on SIP implementations. An advanced tool to construct sophisticated test-cases for SIP is SIPp [11]. SIPp offers the definition of complex and dynamic tests for SIP implementations.

Spam over Internet Telephony (SPIT)

Though not an issue today, it is estimated that Spam over Internet Telephony (SPIT) will become a problem in the future. First, automatic generation of SIP-based phone calls is feasible and cheap. Second, VoIP Spam will be much more intrusive than e-mail Spam is today: A phone will actually ring with each SPIT occurrence (possibly in the middle of the night). VoIP deals with real-time audio signals. Thus, the same countermeasures as for e-mail spam may not work for SPIT. Examples for work in this area are:

[2] Once authentication has taken place, an encryption key for the media stream can - for example - be negotiated with Multimedia Internet Keying (MIKEY) and the media stream can be secured via the Secure Real-Time Transport Protocol (SRTP).

- A comparison to e-mail spam and an overview on possible solutions against SPIT in SIP networks can be found in [12].
- SIP extensions for feedback on SPIT detection and prevention are proposed in [13].
- A prototype for an anti-SPIT solution has been described in [14].

Lawful Interception

Most countries legally allow for authorized wiretapping of telephone calls by law enforcement agencies, so-called Lawful Interception. Depending on the use case and national law, Lawful Interception legislation may apply to VoIP. However, Lawful Interception for VoIP is much harder than in the PSTN due to the following technical facts:

- The SIP provider and the Internet Service Provider (ISP) may be different.
- Signalling and payload usually take a different route, traffic is only linked in terminals.
- The signalling and payload of the conversation may be encrypted.

Thus, in order to reliably deploy Lawful Interception for VoIP it would be necessary to a) intercept all SIP traffic and b) intercept the network traffic in real-time of a provider not known prior to call-setup. Several scientists have realized the potential problems of Lawful Interception for VoIP. They have made a proposal arguing that the benefit of Lawful Interception for VoIP may be outweighed by the negative consequences for society [15].

5. Security Problems for P2P-SIP

Although SIP is specified as a client-server protocol [1], recent proposals suggest using SIP in a peer-to-peer (P2P) setting [16], [17]. This approach of using a peer-to-peer network as a substrate for SIP signalling is frequently called P2P-SIP. P2P-SIP is currently discussed in the IETF and many internet drafts exist (see [18]). The general approach is as follows: Instead of servers, a P2P network is used for SIP registration and user location. Researchers have proposed to use a Distributed Hash Table (DHT) instead of SIP servers for user location and user registration [16], [17]. Distributed Hash Tables have been developed to reliably store and retrieve content in a distributed network[3].

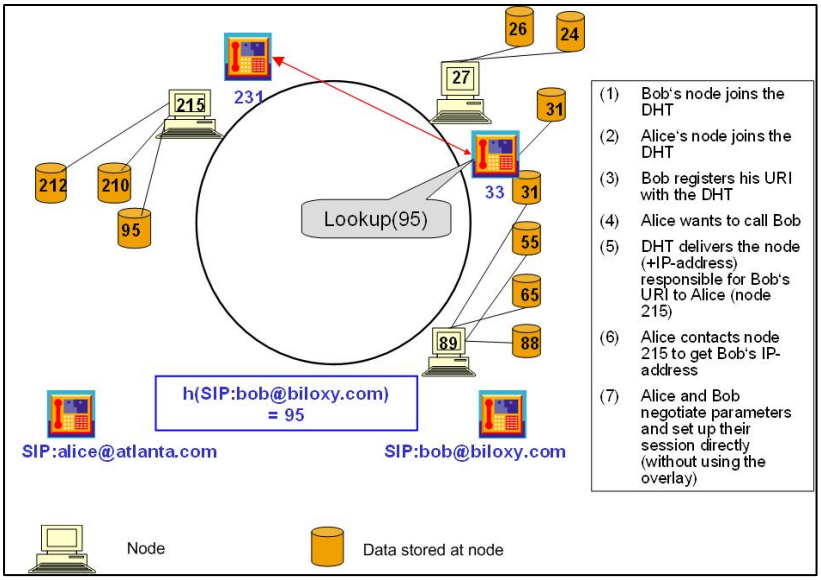

Figure 2 – Simplified overview of locating a SIP user with P2P-SIP

Example of operation

For P2P-SIP, the content stored in the DHT is the binding of user location and SIP-URI. The index to some binding gets computed by hashing the desired SIP-URI. If a node wants to request the current location for a SIP-URI, it inserts a lookup request into the network. This lookup request is routed to the node responsible for the index[3]. That node delivers the content to the requesting node. Figure 2 exemplifies how locating a SIP communication partner is done in a P2P-SIP network [19]: Two users, Alice and Bob, want to communicate. In order to use the network, Alice's and Bob's user agents have to join the network (1), (2). In the example, Alice joins as node 33 and Bob joins as node 231. In order to receive calls, Bob has to register his SIP-URI with the P2P network (3). To do so, Bob hashes his SIP-URI and stores his current location at the node responsible for hash(SIP-URI). In the example, Bob's SIP-URI hashes to 95 and is stored at node 215 in the network. If Alice wants to call Bob (4), she can ask the P2P-network for the node which is responsible for Bob's URI: She computes the index for Bob's URI by hashing his SIP-URI and invokes lookup(index) as a service offered by the network. The lookup request gets routed through the P2P network and finally returns the IP-address and port of the node responsible for the requested content (5). Alice can then contact that node (node 215 in the example) directly to receive Bob's location (6). Finally, Alice can contact Bob (7).

[3] For a tutorial on structured P2P networks and Distributed Hash Tables the reader is referred to [24].

Security Challenges

The P2P paradigm introduces new security threats to SIP. Most important, the lack of a central authority makes authentication of users and nodes difficult. Without authentication, adversary nodes can falsify messages, drop messages, or spoof identity. In this way malicious nodes can launch man-in-the-middle or denial-of-service attacks. With no trusted authority that certifies identities, adversary nodes can control a large fraction of a distributed system [20]. Specific security problems that need to be considered in a P2P-SIP system are (see [21] for a detailed discussion):

- Secure node ID mapping
- Secure routing
- Bootstrapping
- Anonymity
- Identity Enforcement
- Free Riding
- Lawful Interception
- Spam Prevention
- Emergency Services

Options to make P2P-SIP secure

In order to make a P2P network secure for SIP, in general the following approaches can be taken [21]:

1. *Central Authority:* In principle, a trusted authority can certify nodes' identities in the network. This would enable authentication between nodes and prevent most attacks. However, a central authority would need to be trusted and accepted by all users in the system. Furthermore, public key infrastructures do not scale well in practice.
2. *Distributed Solution:* Instead of a central authority, a distributed mechanism could provide authentication in a P2P-SIP network. For instance, reputation management systems assign trust values to nodes in a distributed fashion. However, most reputation management systems that have been developed focus on file-sharing and are therefore not (yet) applicable to P2P-SIP.
3. *Other Approaches:* Because the solutions presented above are not satisfactory in all application scenarios, researchers are trying to develop alternatives. Examples of such alternative approaches that could be applied to P2P-SIP are self-certifying SIP-URIs [19] or a trusted randomness service [22].

6. Conclusion

The intention of this paper has been to present an overview of important challenges and current activities on SIP security. Due to many threats, challenges, and the huge amount of work going on, we were only able to give an overview on some important aspects of SIP security. Many problems for VoIP security have not yet been solved satisfactorily. SIP is used to initiate VoIP communications. Thus, SIP security will remain an active and interesting research area in the near future.

References

[1] J. Rosenberg, H. Schulzrinne et al., "SIP: session initiation protocol", RFC 3261, 2002

[2] J. Posegga, J. Seedorf, "Voice over IP: Unsafe at any Bandwidth?", Eurescom Summit 2005 – Ubiquitous Services and Applications, Heidelberg, April 27-29, 2005, pp. 305-314, VDE Verlag

[3] Voice over IP Security Alliance, "VoIP Security and Privacy Threat Taxonomy", Public Release 1.0, http://www.voipsa.org/Activities/taxonomy.php, Oct. 2005

[4] Bundesamt für Sicherheit in der Informationstechnik, "VOIPSEC Studie", http://www.bsi.bund.de/literat/studien/VoIP/index.htm

[5] P. Zimmermann, A. Johnston, J. Callas, "ZRTP: Extensions to RTP for Diffie-Hellman Key Agreement for SRTP", http://www.philzimmermann.com/docs/draft-zimmermann-avt-zrtp-01.html, internet draft, March 2006

[6] J. Peterson, C. Jennings, "Enhancements for Authenticated Identity Management in the Session Initiation Protocol (SIP)", draft-ietf-sip-identity-06 (work in progress), October 2005.

[7] V. Gurbani, F. Audet, D. Willis, "The SIPSEC Uniform Resource Identifier (URI)", internet draft (work in progress), June 2006

[8] Cisco Security Advisory: Multiple Vulnerabilities in Cisco IP Telephones, http://www.cisco.com/warp/public/707/multiple-ip-phone-vulnerabilities-pub.shtml

[9] PROTOS Test-Suite: c07-sip, http://www.ee.oulu.fi/research/ouspg/protos/testing/c07/sip/, University of Oulu, Finland

[10] sipsak homepage, SIP swiss army knife, http://www.sipsak.org/

[11] SIPp, http://sipp.sourceforge.net/

[12] J. Rosenberg, C. Jennings, "The Session Initiation Protocol (SIP) and Spam", draft-ietf-sipping-spam-03, internet draft (work in progress), October 2006

[13] S. Niccolini, S. Tartarelli, M. Stiemerling, S. Srivastava, "SIP Extensions for SPIT identification" , draft-niccolini-sipping-feedback-spit-02, internet draft (work in progress), August 2006

[14] S. Niccolini, "SPIT and SPIM", http://www.iptel.org/voipsecurity/workshop/program_1stjune2006.php, 3rd VoIP Sec. Workshop, June 2006, Berlin, Germany

[15] S. Bellovin, M. Blaze, et al., "Security Implications of Applying the Communications Assistance to Law Enforcement Act to Voice over IP", http://www.itaa.org/news/docs/CALEAVOIPreport.pdf

[16] K. Singh, H. Schulzrinne, "Peer-to-Peer Internet Telephony using SIP", Proceedings of the international workshop on Network and operating systems support for digital audio and video, Stevenson, Washington, USA, June 2005, pp. 63-68, ACM Press

[17] D.A. Bryan, B.B. Lowekamp, C. Jennings, "SOSIMPLE: A Serverless, Standards-based, P2P SIP Communication System", Proc. of the International Workshop on Advanced Architectures and Algorithms for Internet Delivery and Applications, Orlando, FL, June 2005, IEEE Press

[18] www.p2psip.org

[19] J. Seedorf, "Using Cryptographically Generated SIP-URIs to Protect the Integrity of Content in P2P-SIP", Third Annual VoIP Security Workshop, June 2006, Berlin, Germany, to appear in ACM Digital Library

[20] J. R. Douceur, "The sybil attack", Revised Papers from the First International Workshop on Peer-to-Peer Systems, Cambridge, MA (USA), March 2002, Lecture Notes In Computer Science, Vol. 2429, Springer

[21] J. Seedorf, "Security Challenges for P2P-SIP", IEEE Network Special Issue on Securing Voice over IP, September 2006

[22] T. Condie, V. Kacholia, S. Sankararaman, P. Maniatis, J.M. Hellerstein, "Maelstrom: Churn as Shelter", University of California at Berkeley Technical Report No. UCB/EECS-2005-11, November 2005

[23] B. Goode, "Voice over internet protocol", Proc. of the IEEE, Vol. 90, No. 9, September 2002, pp. 1495-1517

[24] R. Steinmetz, S. Götz, S. Rieche, "Distributed Hash Tables", in P2P Systems and Applications, R. Steinmetz and K. Wehrle (Eds.), LNCS 3485, pp.79-93 & pp. 95-117, 2005