SECURITY CONSIDERATIONS IN CURRENT VOIP PROTOCOLS

Steffen Fries
Corporate Technology, Siemens AG, Otto-Hahn-Ring 6, 81730 Munich, Germany

Keywords: Voice over IP, security, encryption, authentication, integrity, SIP, H.323, Megaco, MGCP, SRTP.

Abstract: This document describes current state of the art security functionality provided in the four mainly used and standardized Voice over IP (VoIP) signaling protocols, as there are the Session Initiation Protocol (SIP), H.323, Megaco, and the Media Gateway Control Protocol (MGCP). It outlines the security provided by the protocols itself or by dedicated security extensions including lower layer security protocols like Transport Layer Security (TLS) or IPSec. Moreover, vulnerabilities, which still remain in protocols or certain scenarios, are depicted as well. Furthermore discussed are also security approaches for the media data provided by the Secure Real-time Transport Protocol (SRTP) and associated key management schemes. Conclusions are given by identifying work areas, in which further security related work in the area of multimedia communication in general and VoIP in specific has to be done.

1 INTRODUCTION

Voice over IP is one of the driving factors for convergent networks. It targets the migration from current circuit switched networks to packet based networks for voice communication. VoIP is already being used in enterprise and carrier environments for VoIP trunking to connect legacy Private Branch Exchange (PBX) via IP as well as for direct user connectivity. Moreover VoIP is offered by public service providers for residential customers.

Critical requirements on information security cannot be satisfied by relying on "trust by wire security" as done within traditional telephone network architectures. The distributed and heterogeneous VoIP system architecture itself enables attacks against integrity and confidentiality of data communicated over the packet network. Inevitable threats to the availability of the involved components are evident. Therefore comprehensive and consistent security architectures are necessary prerequisites for running VoIP communication systems based on bundled multimedia standards. Note that in VoIP scenarios signaling and media may traverse different communication path and thus pose another challenge to the overall security approach.

VoIP communication protocols and associated security is currently being standardized mainly in four standardization organizations. While the IETF (Internet Engineering Task Force) and ITU-T (International Telecommunication Union) define protocols for VoIP like SIP, SRTP (IETF), and H.323 (ITU-T) as well as surrounding protocols, ETSI and 3GPP (Third generation Partnership Project) define the architecture for NGN (Next Generation Network) utilizing these protocols.

Typical threats to information security in general and to VoIP in specific are eavesdropping or wiretapping, misuse of service (e.g., through masquerading), and manipulation.

An overview about information security mechanisms in the commonly used signaling and media protocols is provided as well as recently defined security extensions to cover further use cases. It will be shown that security mechanisms are already incorporated into the main VoIP protocols covering a wide extend of known vulnerabilities.

Challenges for further security related work are misuse of the multimedia protocol's functionality for SPIT (SPAM over Internet Telephony), Denial of Service attacks (DoS). Moreover, besides basic call more advanced features are going to be supported in communication scenarios leading to further security requirements to be met.

2 BASIC SCENARIOS AND DEPLOYMENTS

VoIP is gaining more momentum and will be offered in enterprises and public carriers for business and

Fries S. (2006).
SECURITY CONSIDERATIONS IN CURRENT VOIP PROTOCOLS.
In *Proceedings of the International Conference on Security and Cryptography*, pages 183-191
DOI: 10.5220/0002098601830191

also personal use. The deployment of VoIP can be done using different approaches providing also a possible migration strategy.

A first step for introducing VoIP services is often the trunking of voice connections between different PBX's via packet based networks as shown in Figure 1. This option does not require the use of IP enabled clients. Thus, VoIP can be introduced transparent to the end user.

A next step in the deployment is consequently the introduction of VoIP enabled networks in the source and/or the target domain, including the appropriate endpoints. These endpoints are able to communicate completely over the packet based network, while gateways ensure the connectivity to legacy PBX systems.

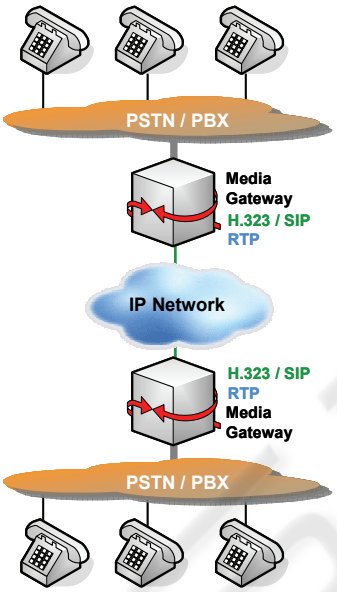

Figure 1: VoIP Trunking.

A general scenario for VoIP is depicted in Figure 2 as an example for larger enterprises or for carrier deployments. In these environments the media data pose a high load on the gateways, when offering connections to the Public Switched Telephone Network (PSTN). Therefore multiple media gateways can be coordinated by a single gateway controller.

The support for multiple gateways, targeting a higher availability rate for connections to the PSTN as well as cost reduction through the use of a single controller for multiple gateways, will be discussed with focus on the security of the supporting protocols in section 5.3.

For smaller enterprises or home offices the media gateways and the media gateway controller collapse to a single device, eliminating also the need for additional gateway control protocols.

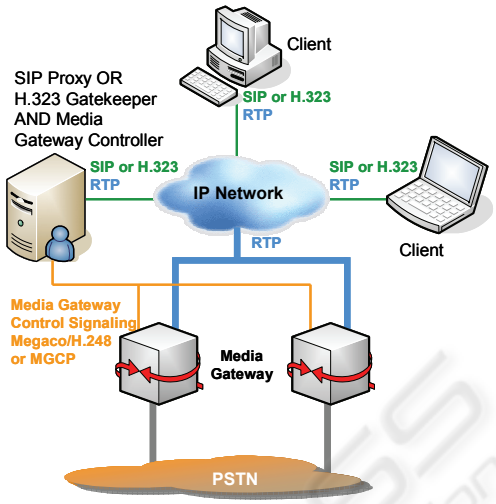

Figure 2: Multiple Gateway Support.

The final stage of VoIP would be a pure packet based voice communication network and may be seen as subset of the scenario depicted above. Because of this and as PSTN-based telephony will exist in parallel to VoIP for quite some years, this scenario is not further considered here.

3 TYPICAL VULNERABILITIES AND THREATS

A threat to information security can be defined as attempt at unauthorized access to an object by an attacker.

Besides the general threats for packet-based IP networks, within VoIP there are some dedicated threat targets:

- **Identity**: Attackers may spoof the identities of service subscribers to misuse services and produce toll fraud. They may also spoof server identities to get access to user related information (one form of spoofing server identities is commonly known as phishing).
- **Privacy**: Eavesdropping or wiretapping of insecure communication of signaling and/or media connections may lead to loss of sensitive information. Examples are PINs or passwords transmitted over the signaling channel or (spoken) personal information transmitted over the media channel.
- **Availability**: Voice or multimedia services are susceptible to Denial of Service attacks, call hijacking or call disruption. As multimedia-services are real-time services, the importance of this property increases. Also, if VoIP is used as a

complete substitution for PSTN based telephony, emergency call handling needs to be provided as it will be required by regulation.

All of these properties are crucial for the general acceptance of VoIP in business and personal communication. Typical resulting security requirements are depicted next.

4 SECURITY REQUIREMENTS

As VoIP is transmitted in (former) data networks, similar security requirements as known for other services in such a network apply. These are in the first place:

- **Authentication**: The property that the claimed identity of an entity is correct.
- **Authorization**: The process of giving someone permission to do or have something.
- **Integrity**: The property that information has not been altered in an unauthorized manner.
- **Confidentiality**: The property that information is not made available or disclosed to unauthorized individuals, entities or processes.

Additional challenges for VoIP security are given through the different communication path for signaling and media on one hand. Especially the dynamic port assignment used for the media connections needs to cope with existing security infrastructure, like NAT devices and Firewalls.

On the other hand media security or, to be more precise, the key management for the media security is crucial (cf. also section 6.3). Here more complex communication scenarios than simple point-to-point connections need to be supported. In VoIP these scenarios comprise the support for early media, i.e., a call initiator will receive media data prior to receiving an answer on the signaling path. Another example is the forking of calls to reach different endpoints simultaneously. These endpoints may belong to the same user or form for instance a working group (commonly known is the 'group pickup feature'). Especially in the latter case user authentication and identity provision is required and may pose potential obstacles.

5 VOIP SIGNALING PROTOCOLS

Currently there exist several protocols for VoIP, some of them are competitive, e.g., SIP and H.323, and some are complementary, e.g., SIP and MGCP. The following subsections comprise the discussion of security focused on state of the art description of the signaling protocols SIP (Rosenberg et al, 2002), H.323 (ITU-T, 2003), Megaco/H.248 (Green, Ramalho and Rosen, 2000), and MGCP (Arango et al, 1999). Note, that the first two protocols provide inherent security means, while the others mainly rely on IPSec.

Common analyses show that the number of SIP deployments increases more compared to H.323.

IPSec in general is applicable for UDP, TCP and SCTP based signaling and may be used to provide authentication, integrity and confidentiality for the transmitted data. It supports end-to-end as well as hop-by-hop scenarios. Thus, it may be used to provide security functions for all of the above protocols. Nevertheless, it may not be applicable straight forward, as it is often not considered by the signaling standard itself and may not provide for advanced communication scenarios. Moreover, due to the message overhead of IPSec it may not be easily applicable for real-time media communication. Therefore alternative approaches have been taken.

5.1 SIP

The Session Initiation Protocol is specified by the IETF in RFC3261 SIP (Rosenberg et al, 2002). It enables the initiation and control of communication sessions, which may be VoIP or complete multimedia sessions. The general architecture is shown in Figure 2, were the server is called SIP proxy or redirect server. The signaling of control information is done using SIP, while media is sent using RTP.

SIP is a text–based Internet protocol similar to protocols like Hypertext Transfer Protocol (HTTP) and Simple Mail Transport Protocol (SMTP) in contrast to other protocols like H.323 basing on the Abstract Syntax Notation (ASN.1). ASN.1 requires additional encoding and decoding operations. Nevertheless, if security using Secure Multipurpose Internet Mail Extensions (S/MIME) is provided, ASN.1 becomes an issue also for SIP. SIP is independent of the transport layer (supports User Datagram Protocol (UDP), Transport Control Protocol (TCP) and Stream Control Transmission Protocol (SCTP)) and has been designed to be not restricted for Voice over IP.

SIP already considers security and provides measures for the most common scenarios. As SIP is flexible with regard to extending the protocol, several security enhancements have already been standardized. Furthermore, as new scenarios arise, new

security extensions are being proposed. This is also depicted in section 5.1.2.

Recent discussions also comprise the usage of SIP to realize peer-to-peer VoIP systems. This will pose new security requirements to VoIP. Main work for security can be seen here in the area of distributed identities, prevention of spoofing, and denial of service.

5.1.1 SIP Inherent Security Features

As stated above RFC3261 provides several security features, which are depicted next.

Signalling data authentication using HTTP Digest Authentication

SIP digest authentication is based on the HTTP digest authentication defined in RFC2617 describing a simple challenge-response paradigm. The remote end is challenged using a nonce value. A valid response contains a checksum (by default, the MD5 checksum) of the user name, the password, the given nonce value, the HTTP method, and the requested URI (Uniform Resource Identifier). In this way, the password is never sent in the clear. Nevertheless, there are some deficiencies in the usage of the HTTP Digest scheme, as it does not provide complete message integrity and may not be applied to all messages. Here TLS kicks in, which can be used to enhance the signaling protection.

Approaches for protecting the signaling data using TLS

RFC3261 mandates the support of TLS for SIP server components (proxies, redirect servers, and registrars) to protect SIP signaling. Using TLS for User Agents (UAs) is recommended. TLS protects SIP signaling messages against loss of integrity, confidentiality and against replay. It provides integrated key-management with mutual authentication. TLS is applicable hop-by-hop between UAs and proxies or between proxies. The SIP-Secure (SIPS) scheme defined in RFC3261 requires the usage of TLS to protect the signaling until the last proxy in the call flow. Drafts are currently being discussed addressing the deficiency regarding the last hop usage of TLS (see section 5.1.2). TLS is usually applied in a hop-to-hop fashion. Note that the calling client does not get a confirmation about the usage of TLS till the final recipient. It merely has to trust the proxies acting according to the standard. The drawback of TLS in SIP scenarios is the requirement of a reliable transport stack (TCP-based SIP signaling). The recent development of Datagram-TLS (DTLS) may provide for using a TLS-like protection also for UDP based signaling. RFC3261 mandates the support of a dedicated TLS crypto scheme, which ensures interoperability. Other schemes may be negotiated as part of the TLS handshake.

S/MIME to protect SIP message body data in an end-to-end fashion

RFC3261 recommends the IETF defined standard S/MIME (RFC3850 and RFC3851) to be used for end-to-end protection of signaling message payloads. S/MIME within SIP supports the following security services:

- Authentication and Integrity Protection of Signaling Data
- Confidentiality of Signalling Data

Note that S/MIME is not widely used in current SIP deployments and therefore not further discussed.

5.1.2 SIP Security Extensions

Meanwhile more complex SIP communication scenarios have been worked out, requiring additional Request For Comments (RFCs) and new drafts ensuring authentication and integrity and also confidentiality. Identity is one of the intensively discussed topics within the SIP community. This can be seen on the variety of documents concerning identity.

In the following an overview about a subset of the most interesting security extensions is given. These extensions solve different problems identified in SIP communication scenarios.

SIP Authenticated Identity Body

SIP Authenticated Identity Body (AIB) defines a generic SIP authentication token. It is provided by adding an S/MIME body to a SIP request or response in order to provide reference integrity over its headers. The AIB is a digitally signed SIP message or message fragment. This approach has been standardized as RFC3893.

Enhancements for the Authenticated Identity Management in SIP

The security options of RFC3261 utilize asymmetric security in terms of certificates and corresponding private keys. The distribution of these credentials is often complex and limited to a dedicated domain. Certificates can be used to provide means for identity management, but this may not be uniquely defined. An example can be drawn by two different administrative domains with domain-specific PKIs. The current draft (Peterson and Jennings, WiP) addresses this limitation by defining an authentication service providing assertions for the user identity (Address of Record) transmitted in the

header of SIP requests. This authentication service is responsible for a dedicated domain.

Certificate Management Service for SIP

As stated before, certificate distribution is often cumbersome and a global Public Key Infrastructure (PKI) does not exist so far. Several security services for SIP relay on certificates, so their distribution becomes a crucial part. The draft (Jennings and Peterson, WiP) describes a solution using a certificate server to provide certificates and even complete user credentials using a Simple Authentication and Security Layer (SASL) like approach. Users have the option to store their complete credentials or only the certificate over a secure connection on a central server. This may be useful, when users are in need to transmit credentials between different devices.

Connection Reuse / Outbound Connections

SIP defines the usage of TLS to protect the signaling. But it requires only server components to support mutual authentication. This leads to the problem that clients without a TLS certificate cannot receive inbound calls over TLS. This is due to the fact that SIP communication is done over distinct ports for inbound and outbound traffic. Furthermore, an endpoint without a certificate TLS may not be run in server-mode. To handle this deficiency two drafts (Jennings and Hawrylyshen, WiP) and (Mahy, WiP) describe the possibility of reusing already established TLS connection, which were initiated by the client. The basic idea is the provision of a flow identifier for the different streams within an established session to identify inbound and outbound traffic.

5.2 H.323

H.323 (ITU-T, 2003) is an umbrella recommendation defined by the ITU-T (International Telecommunication Union). H.323 addresses call control, multimedia management, and bandwidth management as well as interfaces between LANs and other networks. It includes point-to-point and multipoint conferences. The first version has been defined in 1996 and has been improved consistently. Security features have become part of the standard.

The general architecture is similar to SIP and can also be depicted using Figure 2, were the server is called H.323 gatekeeper. The signaling of control information is done using H.225, while media is sent using RTP.

The call establishment can be performed in various ways, as there are gatekeeper routed calls, direct routed calls with gatekeeper (for address resolution) and plain direct routed calls.

The security functionality is defined within H.235 (ITU-T, 2005). H.235 is splitted in 9 subgroups, describing so-called profiles for security in H.323.

5.2.3 H.323 Security Profiles

The security profiles may be distinguished as profiles for signaling integrity and authentication and for media security. They are related to the call models used in H.323. Note that the former H.235 annexes have been reorganized in a new form (H.235.x) recently.

- **H.235.0** provides the framework of the subsequent H.235 standards within H.323.
- **H.235.1** provides signaling integrity and authentication using mutually shared secrets and keyed hashes (HMAC-SHA1-96) in gatekeeper-routed scenarios. This profile is widely implemented in available H.323 solutions.
- **H.235.2** provides signaling integrity and authentication using digital signatures on every message in gatekeeper-routed scenarios. Since signature generation and verification is costly in terms of performance, this profile may not gain momentum.
- **H.235.3** is a hybrid approach using both, H.235.1 and H.235.2. During the first handshake a shared secret establishment is performed, protected by digital signatures. Afterwards keyed hashes are used for integrity protection, based on the established shared secret.
- **H.235.4** is the adaptation of H.235.1 for direct routed call scenarios with gatekeeper, were the gatekeeper provides the key material for securing the direct routed call.
- **H.235.5** specifies a framework for secure mutual authentication during the registration and admission phase using weak shared secrets in combination with Diffie-Hellman key agreement for stronger authentication during call signaling. Extensions to the framework to permit simultaneous negotiation of TLS parameters for protection of a subsequent call signaling channel are also provided.
- **H.235.6** describes the voice encryption profile. This profile relies on certain security services as part of the call signaling and connection setup procedures; e.g., the Diffie-Hellman key agreement and other key management functions and is not compatible to SRTP.
- **H.235.7** can be seen as evolvement of H.235.6 as it provides the framework for providing key ma-

terial for media encryption based on MIKEY and SRTP within H.235.
- **H.235.8** describes another approach for key management for SRTP as 'Key Exchange for SRTP using secure signaling channels'. This is a sdescriptions-like (cf. section 6.3.2) approach were all parameter necessary for voice encryption are sent in-band protected by underlying security (e.g., like TLS) or by using Cryptographic Message Syntax (CMS).
- **H.235.9** depicts the most recent extension of H.235; the security gateway support for H.323. The document defines a method for the discovery of security gateways in the signaling path between communicating H.323 entities, and for sharing of security information between a gatekeeper and a security gateway in order to preserve signaling integrity and privacy when crossing network boundaries.

H.235 security features are defined to complete H.323 scenarios. Additionally, the interworking with security features of other multimedia protocol suites, e.g., SIP, is considered. An example is the key management and media data encryption using H.235.7 and H.235.8.

5.3 Gateway Decomposition

Gateway decomposition describes the separation of signaling and media functionality. The general goal of gateway decomposition is a cost reduction through the use of only one media gateway controller responsible for several media gateways. The media gateway controller connects to common multimedia signaling protocols like SIP or H.323, described above and controls the media gateways via a trivial protocol. The general architecture approach is shown in Figure 2.

There are two commonly used protocols within the context of gateway decomposition – Megaco (Green, Ramalho and Rosen, 2000) and MGCP (Arango et al, 1999). Megaco/H.248 has basically the same architecture as MGCP. Also the commands are similar. However, the protocol models are quite different.

Generally, these gateway control protocols provide complementary functionality and architectural components to plain SIP or H.323 architectures.

5.3.1 Megaco

In June 1999 IETF Megaco Working Group (WG) and ITU-T provided a unified document describing a standard protocol for interfacing between Media Gateway Controllers (MGCs) and Media Gateways (MGs) Megaco/H.248. It is expected to win wide industry acceptance as the official standard for decomposed gateway architectures.

RFC2805 recommends security mechanisms that may be in underlying transport mechanisms, such as IPSec. H.248 goes even a step further by requiring that IPSec shall be implemented, where the underlying operating system and the transport network support IPsec. Implementations of the protocol using IPv4 shall implement the interim AH scheme.

The interim AH scheme is the usage of an optional AH header, which is defined in the H.248 protocol header. The header fields are exactly those of the SPI, SEQUENCE NUMBER and DATA fields as defined in RFC4302. The semantics of the header fields are compliant with the "transport mode" of RFC4302, except for the calculation of the Integrity Check Value. The interim Authentication Header (AH) scheme does not provide protection against eavesdropping and replay attacks (the sequence number in the AH may overrun when using manual key management since re-keying is not possible).

5.3.2 MGCP

MGCP is currently being maintained by PacketCable™ and the Softswitch Consortium™. In October 1999, MGCP was finally converted into an informational RFC2705.

Regarding security, there are no mechanisms designed into the MGCP protocol itself. The informational RFC2705 (Arango et al, 1999) refers to the use of IPSec (either AH or Encapsulated Security Payload (ESP)) to protect MGCP messages. Without this protection an adversary could setup unauthorized calls or interfere with ongoing authorized calls.

Beside the usage of IPSec, MGCP allows the call agent to provide gateways with session keys that can be used to encrypt the payload of the Real-time Transport Protocol (RTP), protecting against eavesdropping. To achieve this RTP encryption, described in RFC3550 (Schulzrinne, 2003) may be applied. Session keys can be transferred between the call agent and the gateway by using SDP Handley and Jacobson, 1998) either directly or using dedicated key management extensions.

6 MEDIA DATA SECURITY

The signalling protocols described above do not consider the encryption of media data directly. Nevertheless, they allow the negotiation of security

features and also the key management for the security associations of the media channels.

Real-time media data in VoIP environments is usually transmitted using RTP. For RTP basically two approaches for security provisioning exist, which are discussed in the following.

6.1 RTP/RTCP

RFC3550 defines the RTP protocol as well as the associated Real Time Control Protocol (RTCP). Using the optional RTP encryption as defined in RFC3550 provides for confidentiality for media data using the Data Encryption Standard (DES) as default method. The functionality is limited as the following issues are not considered in the RTP standard:

- key management
- authentication and message integrity (not feasible without a key management infrastructure)
- replay protection

Because of these points, the usage of this option is not widely deployed.

6.2 SRTP/SRTCP

The RTP standard provides the flexibility to adapt to application specific requirements with the option to define profiles in companion documents. SRTP (Baugher et al, 2004) has been recently defined as RTP profile and addresses the limitations. This profile provides confidentiality using the Advanced Encryption Standard (AES, the successor of the DES), message authentication (using keyed hashes) and replay protection to the RTP/RTCP traffic.

SRTP does not define an own key management and relies on solutions like MIKEY or sdescriptions (cf. section 6.3). The key management messages have to be transported by the signaling protocol.

For SDP based protocols like SIP or MGCP the key transport can be transported as part of SDP itself, while for H.323 extensions to the base protocol are defined through H.235.7 and H.235.8 as described above.

6.3 Key Management

In contrast to common data applications, the key management for negotiating the VoIP media security has to cope with several additional requirements, as there the real-time conditions, support of advanced communication scenarios like forking, retargeting and also conferencing. Moreover, due to the nature of VoIP to separate signaling and media traffic, the key management may be done along the path of the signaling data, the media data, or combined.

Currently there is a broad discussion about the key management for SRTP using SIP within the IETF as meanwhile 13 different approaches exist. They utilize different cryptographic techniques like pre-shared keys, Diffie-Hellman key agreement, digital signatures, or asymmetric encryption (cf. (Audet and Wing, WiP)). Due to the different mechanisms used, the approaches are not interoperable. Moreover, none of the current approaches provides a solution for all of the scenarios stated above.

In the following subsection two promising key management approaches out of the 13 are depicted in more detail. Both belong to the cluster of signaling path key management approaches and have been deployed already. Examples for media path key management are given through ZRTP (cf. (Zimmermann, WiP)) and DTLS usage (cf. (McGrew and Rescorla, WiP)).

6.3.1 Multimedia Internet Keying

MIKEY (Multimedia Internet Keying) has been defined in the IETF within RFC3830 (Akko, 2004). It defines an authentication and key management framework that can be used for real-time applications (both for peer-to-peer communication and group communication). In particular, RFC3830 is defined in a way to support SRTP in the first place but is open to enhancements to be used for other purposes too. MIKEY has been designed to meet the requirements of initiation of secure multimedia sessions. Such requirements are for instance the establishment of the security parameters for the multimedia protocol within one round trip.

Another requirement is the provision of end-to-end keying material, and also independence from any specific security functionality of the underlying transport layers.

MIKEY defines several options for the user authentication and negotiation of the master keys with a maximum of a complete round trip as there are:

- Symmetric key distribution (based on pre-shared keys and keyed hashes), which may proceed with a single message.
- Asymmetric key distribution (based on asymmetric encryption) may proceed with a single message.
- Diffie Hellman key agreement protected by digital signatures (complete roundtrip)

Unprotected key distribution, i.e., without authentication, integrity, or encryption, is also possi-

ble, but not recommended without any underlying security like TLS or similar. This use case is comparable with the sdescription approach described below.

Two MIKEY enhancements exist as drafts, which are likely to advance to a RFC soon.

- Diffie Hellman key agreement protected by symmetric pre-shared keys and keyed hashes
- Asymmetric (encrypted) key distribution with in-band certificate provision

The default and mandatory key transport encryption is the AES in counter mode, where MIKEY references RFC3711. MIKEY uses a 160-bit authentication tag, generated by HMAC with SHA-1 as mandatory algorithm described in the associated RFC2104. Also mandatory, when asymmetric mechanisms are used, is the support of X.509v3 certificates for public key encryption and digital signatures.

Recently the usage of elliptic curves has been proposed targeting performance saving and enabling the use of shorter cryptographic key material by keeping the same level of security compared to the currently used RSA.

6.3.2 Security Descriptions

Besides MIKEY a second key management for SRTP has been proposed in the IETF, which utilizes the plain Session Description Protocol (SDP). The approach is based on the offer answer model of SDP and transmits all necessary SRTP parameter in a new attribute field and is called sdescriptions (cf. (Flemming and Baugher, WiP)).

The protection of this field is left to SIP itself (applying S/MIME in an end-to-end fashion) or may be done using TLS (in a hop-to-hop fashion) or even IPSec. It therefore nicely integrates with SIP. For other signaling protocols like H.323 there exist similar approaches (through H.235.8). Because of the signaling protocol dependent approach, this solution lacks the support of end-to-end security.

7 CONCLUSION

As shown in this paper, current multimedia protocols already consider security to certain extends.

SIP and H.323 are here the most advanced protocols, as they provide inherent security measures for user authentication and message integrity. Confidentiality can be achieved by additional measures like S/MIME or underlying security protocols TLS or IPSec. Both signaling protocols also provide options to transport a key management for SRTP. Several key management approaches have been proposed leading to the requirement for profiles to ensure easy (inter)operation.

SIP and H.323 suffer from dynamic port signaling as part of the protocol payload, which may lead to problems with widely deployed NAT devices, when message integrity or confidentiality is desired. Within the IETF efforts have been spent to cope with this problem by providing a methodology for Interactive Connectivity Establishment (ICE, (Rosenberg, WiP)).

Security for gateway control protocols is mainly provided by using IPSec as underlying security protocol. Here the communication association is merely between the controller and the associated gateways. Thus, the signaling scenarios are rather simple compared to SIP and H.323.

As the scenarios, in which VoIP is about to be deployed, are getting more complex through the integration of already available features of the legacy telephone system, new security requirements arise. An example is the handling of security properties like user authentication or media session confidentiality in case of call transfer. This is especially becoming interesting in scenarios were the participants of a call cannot rely on the same security infrastructure. Here a global PKI solution could support user authentication. As this is not available right now, alternative approaches are necessary.

Potential obstacles for VoIP usage may arise through the possibility of misusing the infrastructure resulting in Denial of Service. SPIT will pose another potential obstacle. As seen for today's email communication SPAM poses a severe problem for communication. To counter similar threats in VoIP certain measures have to be taken and are already being discussed.

Further challenges are given through the user's request for interoperability, whereby this relates to several facets, like the implementation of a standard through different vendors and also proprietary enhancements. Moreover other solutions are available, which are not standardized so far. While the first point may be covered by regular interoperability tests, the second leads to a subset of common functionality, which ensures interworking. For SIP the definition of this common set of functions is done in the SIP Forum. A prominent example for the third point is Skype, providing security of signaling and media traffic in a proprietary way, not interoperable to SIP or H.323 based clients. Thus security interworking in an end-to-end fashion may not be possible.

Last but not least, security functions provided by signaling and media protocols are only one part of the multimedia puzzle. There is more work to be done in defining the suitable system architecture and deploying multimedia systems in a secure way. Most important, multimedia systems have to be designed into already existing networks under consideration of security, availability, quality of service, and also legal terms. This is also outlined in (Kuhn, Walsh and Fries, 2005). Especially for emergency services appropriate measures have to be taken to meet the balance between security and availability.

ACKNOWLEDGEMENTS

The author would like to thank Wolfgang Klasen and Michael Montag for their valuable review comments and discussions.

REFERENCES

Rosenberg, J., Schulzrinne, H., Camarillo, G., Johnston, A., Peterson, J., Sparks, R., Handley, M. and Schooler, E., 2002, *RFC3261: SIP: Session Initiation Protocol*

ITU-T, 2003, *H.323v5: Packet-based multimedia communications systems*

Rosenberg, J., Work in Progress, *ICE: A Methodology for Network Address Translator (NAT) Traversal for Offer/Answer Protocols,*

Greene, N., Ramalho, M. and Rosen, B., 2000, *RFC2805: Media Gateway Control Protocol Architecture and Requirement,*

Arango, M., Dugan, A., Elliott, I., Huitema, C. and Pickett, S., 1999, *RFC2705: Media Gateway Control Protocol Version 1.0*

Handley, M. and Jacobson, V., 1998, *RFC2327: SDP: Session Description Protocol*

Schulzrinne, H., Casner, S., Frederick, R. and Jacobson, V., 2003, *RFC3550: RTP: A Transport Protocol for Real-Time Applications*

Baugher, M., McGrew, D., Naslund, M., Carrara, E. and Norrman, K., 2004, *RFC3711: The Secure Real-time Transport Protocol*

Audet, F. and Wing, D., Work in Progress, *Evaluation of SRTP Keying with SIP*

Arkko, J., Carrara, E., Lindholm, F., Naslund, M. and Norrman, K., 2004, *RFC3830: MIKEY: Multimedia Internet KEYing*

ITU-T, 2005, *H.235.0: Security framework for H-series*

Arkko, J., Carrara, E., Lindholm, F., Naslund, M. and Norrman, K., Work in Progress, *Key Management Extensions for Session Description Protocol (SDP) and Real Time Streaming Protocol (RTSP)*

Peterson, J. and Jennings, C., Work in Progress, *Enhancements for Authenticated Identity Management in the Session Initiation Protocol*

Jennings, C. and Peterson, J., Work in Progress, *Certificate Management Service for The Session Initiation Protocol,*

Jennings, C. and Hawrylyshen, A., Work in Progress, *SIP Conventions for UAs with Outbound Only Connection,*

Mahy, R., Work in Progress, *Connection Reuse in the Session Initiation Protocol (SIP)*

Flemming, A., Baugher, M. and Wing, D., Work in Progress, *Session Description Protocol Security Descriptions for Media Streams*

Zimmermann, P., Work in Progress, *ZRTP: Extensions to RTP for Diffie-Hellman Key Agreement for SRTP*

McGrew, D. and Rescorla, E., Work in Progress, *Datagram Transport Layer Security Extension to Establish Keys for Secure Real-time Transport Protocol*

Kuhn, D.R., Walsh, T.J. and Fries, S., 2005, *Security Considerations for Voice over IP Systems, SP800-58, US NIST*