STUDIA INFORMATICA 2008
Volume 29 Number 4B (81)

Piotr ZAWADZKI
Silesian University of Technology, Institute of Electronics

THE VOIP COMMUNICATION SECURITY PROTOCOLS

Summary. The presently offered VoIP services pose a serious security problem – they are vulnerable to eavesdropping, impersonation, session hijacking and denial of service attacks. The paper presents security analysis of the proposed VoIP protocol stack, including signaling protocol SIP, key management protocols ZRTP and MICKEY and transport layer security protocol SRTP. The VoIP network security subsystem upgrade path is also described.

Keywords: VoIP, security, protocols

PROTOKOŁY OCHRONY KOMUNIKACJI VOIP

Streszczenie. Powszechnie dostępne usługi VoIP nie zawierają mechanizmów chroniących transmisję, w związku z czym są podatne na podsłuch, przechwytywanie sesji oraz blokadę usługi. W artykule przedstawiono analizę bezpieczeństwa protokołów VoIP z uwzględnieniem protokołu sygnalizacyjnego SIP, protokołów zarządzania kluczami ZRTP i MICKEY oraz protokołu transportowego SRTP. Zaprezentowano również propozycję poprawy proponowanych obecnie sposobów ochrony komunikacji.

Słowa kluczowe: VoIP, bezpieczeństwo, protokoły

1. Introduction

The network and service convergence is the Holy Grail of the telecommunication business. There have been proposed many technologies like Integrated Service Data Network (ISDN) or Asynchronous Transfer Mode (ATM) promising such integration. Recently there has been a lot of excitement around utilization of the IP network as a carrier for voice and video applications due to large improvements in quality of service making real-time communication possible. Voice over IP (VoIP) becomes rapidly a mature technology providing audio quality and bandwidth usage at acceptable levels. Many individuals and organizations are presently considering switching to VoIP technology for many reasons, the cost reduction is one of the most prevailing. The substantial cost reduction, easy installation and the existence of gateways to the Public Switched Telephony Network (PSTN) has already given the impulse to the development of that market.

2. Threats

Switching to VoIP technology has hidden costs related to confidentiality of the calls. The illegal wiretapping in PSTN network is possible only on the so called last mile of the connection and requires physical access to the wires. In contrary, the Internet should be regarded as shared access medium from the security point of view. Unfortunately, security was outside the interest of the designers of the first generation VoIP networks. The present public VoIP installations do not include any security measures and private enterprise solutions usually use Virtual Private Network (VPN) solutions to protect VoIP traffic over the Internet. In the current state of evolution VoIP networks are exposed to many security threats: eavesdropping, impersonation and toll fraud, session hijacking, voip spam [1]. Eavesdropping exploits the lack of confidentiality in the protocols responsible for transporting media streams. The sniffer located at any router across the call path or using Address Resolution Protocol (ARP) poisoning to redirect packet traffic is able to decode transmission and gain the access to the call contents. Such an attack is especially dangerous if unconscious VoIP user uses his phone to access some sensitive information, for instance, to manage a bank account, and enters secret PIN code as DTMF sequence. Very weak authentication methods proposed in VoIP protocols open the possibility to mount the impersonation attack. Presently only two methods are widely used: open password transfer or challenge-response protocol based on MD5 hash function. The first method reveals the access password to any sniffer located between the user and VoIP server. The second version is better, however, it is vulnerable to the dictionary attack. Impersonation of the user opens many possibilities to malicious attackers, and particularly opportunities to make calls on the victim's account.

3. VoIP architecture

The VoIP networks design follow the same pattern, independent of the used protocols stack. The terms used later in this chapter are taken from the Session Initiation Protocol (SIP) description as this is presently the dominant protocol for providing VoIP services. Two main entities may be identified in the network structure [2] (Fig. 1).

1. User Agents are the end devices of the network. They may be realized as the hardphones, softphones or gateways to the PSTN services. User agents are logically divided on User Agent Client (UAC) and User Agent Server (UAS). UAC is responsible for initiating calls and UAS for handling incoming call requests.
2. VoIP servers are responsible for call signaling and optionally for media format conversions. Basic functions of the VoIP server (sometimes delegated to separated units) include registrations of users and redirection of the call requests. The associations between the user name and the IP address of the listening UAS are created during registration and managed by the location server. Next, this information is used to establish a call on behalf of calling party or to inform the caller about the location of the callee's UAS.

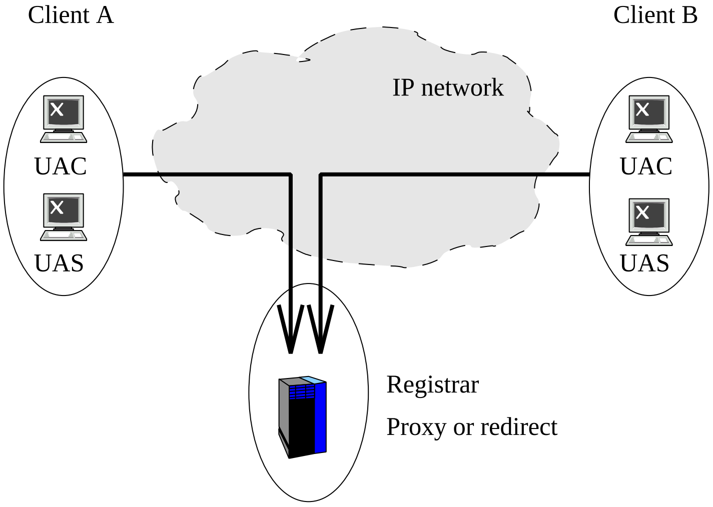

Fig. 1. The components of the SIP network
Rys. 1. Elementy sieci SIP

The VoIP architecture proposed by the IETF is based on the cooperation of different protocols. The proposed protocol stack may be divided into three layers: signaling, session description and media transport. Signaling is the application layer realized by the SIP protocol providing mechanism for creating and terminating VoIP sessions [3]. Session Description Protocol (SDP) encapsulated in SIP messages is responsible for the announcement of the endpoint capabilities and the negotiation of the media coding format used during the call. The Real-time Transport Protocol (RTP) [4] provides conversion of the media streams generated at both ends into the packets and its delivery to the communicating peers.

The SIP uses Universal Resource Identifier (URI) with the syntax similar to email address to locate users. The domain part of the URI usually denotes the DNS name of the registration server, and the user part is the account name assigned to the VoIP service. Sometimes ordinary telephone numbers are assigned to the registered users to make them callable from the ordinary phones equipped only with a numeric dial pad.

The flow of the messages related to the typical call is presented on Fig. 2. Only INVITE message require user authentication, none of those messages is integrity protected, moreover, RTP protocols carries unencrypted codec data. These features open a plethora of possibilities for malicious attackers:

- the weak INVITE authentication – user impersonation,
- the lack of integrity protection – session hijacking using spoofed mid call signalling messages and unauthorized session termination,
- the lack of encryption in RTP messages – call eavesdropping.

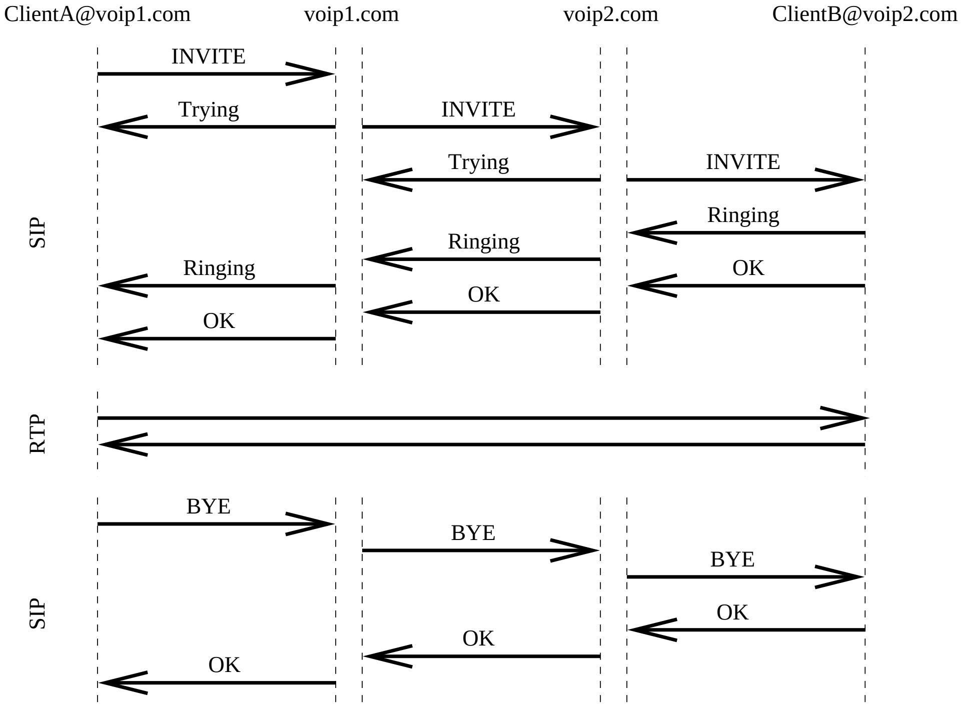

Fig. 2. The setup and termination of the VoIP call
Rys. 2. Zestawienie i zakończenie połączenia VoIP

4. Security improvements

The security threats for the VoIP communication are very similar to those observed in the wireless networks as the Internet should be treated as the medium with shared access form the security perspective. Protection of the networks with the shared medium is a challenging task. The security subsystem should provide:

- strong user to network and network to user authentication,

- pseudorandom session key agreement,
- confidentiality and integrity of the data traffic,
- at least integrity of the network management traffic.

The provision of security for VoIP network is addressed in a series of RFC documents and drafts published by the IETF. The upgrade path proposed for the SIP and RTP protocols is based on the inclusion of the new protocols backward compatible with its predecessors. The goal of the Secure RTP (SRTP) [5], successor of the RTP, is the provision of the media confidentiality and integrity. It is assumed that master session keys are established earlier by some other means, for instance by the described later MIKEY protocol. The master session key and a random number are used to generate the session key utilized for protection of the transported data. The session key serves as the salt in the initialization of the encryption and integrity protection algorithms. The stream generator built around well known Advanced Encryption Standard (AES) block cipher realizes encryption. The stream cipher is separately initialized for each SRTP packet. The packed keys used for confidentiality and integrity protection are composed from the counter included in each frame and the session key. Each packet is integrity protected by the keyed Hash Message Authentication Code (HMAC) based on the standard Secure Hash Algorithm (SHA-1) function. The estimated protection level of the SRTP is very good provided that the keys used for algorithms initialization are properly established.

User and network authentication is provided by the Multimedia Internet KEYing (MIKEY) [6] protocol tunneled in the SDP messages. MIKEY provides authentication and key agreement (AKA) based on preshared secret, public key or authenticated Diffie-Hellman protocol. The initiator is responsible for the selection of the random session key and its secure transport to the responder. In the first mode the session key is encrypted with the preshared secret and the message is integrity protected with keyed HMAC checksum. The initiator may also request a response which is integrity protected with the key recovered from the previous message. The correctness of that checksum proves the second party knows the preshared key. Unfortunately, this method of authentication and key agreement is susceptible to the dictionary attack. The correct guess of the preshared secret enables the user impersonation and an access to the contents of the future and past transmissions. The public key mode and authenticated Diffie-Hellman mode require Public Key Infrastructure (PKI). These protocols assume that each client and proxy server are equipped with the public keys certified by some Certification Authority (CA) known to all parties. The AKA modes based on public keys provide resistance against dictionary and man in the middle (MITM) attacks.

Setting up and managing the PKI infrastructure is a daunting task as it opens a lot of problems related to issuing, revoking and renewal of certificates and their installing on the end devices. The so called blended Diffie-Hellman authentication and key agreement is proposed in the ZRTP protocol [7] performing session key negotiation exclusively for the SRTP protocol. The proposed method is a mix of a preshared key approach with the unauthenticated Diffie-Hellmann key agreement. The Diffie-Hellmann protocol provides resistance against sniffing, but it is vulnerable to the MITM attack. One of improvements of original Diffie-Hellman protocol is based on the introduction of digital signatures to the passed messages. However, digital signatures require working PKI infrastructure but the key assumption of the ZRTP design was to remove the need of PKI installation. The proposed solution uses a Short Authentication String (SAS) for mutual authentication. The value of SAS should be agreed upon by the peers prior to any communication. However, the agreement of SAS on the phones without GUI may be difficult and the ZRTP protocol provides a bootstrap procedure to agree on shared secret with the unauthenticated Diffie-Hellmann protocol. The cached value of the shared secret is used for mutual peers authentication in the subsequent sessions. Unfortunately, the unauthenticated bootstrap procedure is executed also when one of the peers forgets the shared secret because of the reboot, reinstallation or restoring from backup. The protocol recommends issuing the user warning in such a case, but that may be very difficult to realize in the VoIP hardphones or Analog Telephone Adapters (ATA). The protocol leaves the decision to the implementators if the communication should continue upon the detection of the shared secret mismatch. This opens the way to mount the MITM attack.

5. Proposed solution

The standardized solutions are not adequate for carriers of VoIP services. It seems that security architecture similar to that used for Wireless Local Area Network (WLAN) would provide a robust trust transfer between communicating endpoints. The VoIP provider would be obliged to install the certificates issued by a well known CA on its servers and the manufactures of the VoIP equipment should preinstall certificates of those authorities and/or provide an interface to install an additional certificate. The dual layer or tunneled authentication may be then performed (Fig. 3).

- The user authenticates the network. The certificate based network or VoIP server authentication is based on Transport Layer Security (TLS) protocol. On successful accomplishment of this step the user is assured about network identity. From the network point of view user is anonymous and has to be authenticated.
- The network authenticates the user. The TLS protocol provides a confidential communication channel. The messages of user authentication protocol are not accessible

to the sniffer so any, even insecure, methods may be used. The tunnel provides resistance to MITM attack and confidentiality of the subsequent messages.

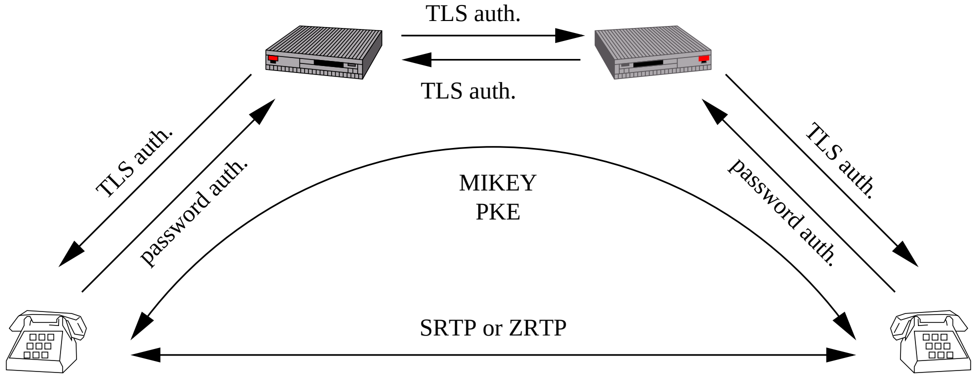

Fig. 3. The SIP security trapezoid
Rys. 3. Trapezoid bezpieczeństwa dla protokołu SIP

The call setup procedure would trigger the authenticated Diffie-Hellmann key agreement protocol executed by the proxy servers and the session key would be securely delivered to the communicating peers. Subsequently nodes may use any secure VoIP traffic protocol for the voice data protection.

6. Discussion

The building of secure VoIP end-to-end communication requires the setup of the trust relation between endpoints. The proposed trust transfers are based on shared secret possession or referring to the arbiter such as trusted Certifying Authority (CA). Unfortunately, both proposed solutions are very difficult to manage. The preshared key scheme cannot be realized because it is impossible to support each end point with the distinct key for each possible end point. The scheme based on certificates serves as the network for shared key transfer. However, the proposed solution requires an issue of the private keys and the matching certified public keys for the peers. The management of PKI is very cumbersome, especially when it comes to revoking certificates, checking and managing Certificate Revocation Lists (CRL). The problems related to the PKI management are well recognized as they are prohibitive to its common installation [8, 9]. These problems are the impulse for the development of the public key architectures without the key certification requirement. Unfortunately, such protocols do not provide a secure trust relation transfer and are susceptible to the MITM attack.

The ZRTP protocol is suitable for VoIP networks, provided that network members use clients with sufficient GUI capabilities. The VoIP solutions may be adequately secured by building dedicated Virtual Private Networks (VPN). However, such solutions are possible only for internal enterprise VoIP services.

The standardized security architectures for the VoIP networks are not mature and require a lot of development effort before the adoption on a large scale. The IETF proposals of the SIP and RTP protocols improvements are difficult to manage and/or vulnerable to well known cryptographic attacks. The presently offered VoIP services pose a serious security problem – they are vulnerable to eavesdropping, impersonation, session hijacking and denial of service attacks.

REFERENCES

1. Butcher D., Li X., Guo J.: Security Challenge and Defense in VoIP Infrastructures. IEEE Transactions on Systems, Man and Cybernetics – Part C: Applications and Reviews, Vol. 37, No. 6, 2007, p. 1152÷1162.
2. Bromirski M.: Telefonia VoIP. BTC, Warszawa 2006.
3. Rosenberg J., Schulzrinne H., Camarillo G., Johnston A., Peterson J., Sparks R., Handley M., Schooler E.: SIP: Session Initiation Protocol. RFC 3261. [@:] http://www.faqs.org/rfcs-/rfc3261.html .
4. Schulzrinne H., Casner S., Frederick R., Jacobson V.: RTP: A Transport Protocol for Real-Time Applications. RFC 3550. [@:] http://www.faqs.org/rfcs/rfc3550.html .
5. Baugher M., McGrew D., Naslund M., Carrara E., Norrman K.: The Secure Real-time Transport Protocol (SRTP). RFC 3711. [@:] http://www.faqs.org/rfcs/rfc3711.html .
6. Arkko J., Carrara E., Lindholm F., Naslund M., Norrman K.: Baugher M., McGrew D., Naslund M., Carrara E., Norrman K.: MIKEY: Multimedia Internet KEYing. RFC 3830. [@:] http://www.faqs.org/rfcs/rfc3830.html .
7. Zimmermann P., Johnston A., Callas J.: ZRTP: Media Path Key Agreement for Secure RTP. RFC draft. [@:] http://tools.ietf.org/draft/draft-zimmermann-avt-zrtp/ .
8. Hunter B.: Simplifying PKI Usage through a Client-Server Architecture and Dynamic Propagation of Certificate Paths and Repository Addresses, Proceedings of the 13th International Workshop on Database and Expert Systems Applications (DEXA 02), 2002, p. 505.
9. Slagell A., Bonilla R., Yurcik W.: A survey of PKI components and scalability issues, Proceedings of IEEE International Performance Computing and Communications Conference, 2006, p. 64.

Recenzent: Dr inż. Bartłomiej Zieliński

Wpłynęło do Redakcji 25 września 2008 r.

Omówienie

Obserwowany w ostatnich latach gwałtowny rozwój Internetu umożliwił wykorzystanie sieci IP do realizacji usług telekomunikacyjnych. Mimo istnienia konkurencyjnych protokołów struktura sieci VoIP jest zawsze taka sama – możemy w niej wyróżnić agentów końcowych odpowiedzialnych za inicjowanie odbierania połączeń oraz serwery obsługujące nomadyczny charakter przyłączeń do sieci i pośredniczących w zestawianiu połączeń (rys. 1). W pierwszych standardach definiujących metody komunikacji VoIP prawie całkowicie pominięto aspekty bezpieczeństwa sieci, koncentrując się na prostocie implementacji i dostępności usług [1, 2, 3, 4]. Obecnie oferowane usługi VoIP podatne są na wiele różnorodnych ataków: kradzież tożsamości, przechwytywanie sesji, podsłuch oraz blokadę usługi. W odpowiedzi w ramach IETF podjęto prace nad protokołami zapewniającymi poufność i spójność ruchu oraz silne wzajemne uwierzytelnianie użytkowników i sieci. Protokół SRTP [5] umożliwia silną kryptograficzną ochronę treści rozmowy, jednak wymaga wcześniejszego uzgodnienia kluczy kryptograficznych przez komunikujące się jednostki. Zarządzanie kluczami kryptograficznymi na etapie zestawienia połączenia umożliwia protokół MIKEY [6], natomiast uzgodnienie klucza bezpośrednio pomiędzy komunikującymi się jednostkami zapewnia protokół ZRTP [7]. Podstawowym problemem obu propozycji jest wymaganie istnienia funkcjonującej infrastruktury PKI dla elementów końcowych sieci w celu ochrony przed atakiem MITM. Założenie takie jest nierealistyczne i prowadzi do dużej komplikacji urządzeń końcowych [8, 9]. W artykule zaproponowano architekturę (rys. 3), w której wymagane jest funkcjonowanie PKI jedynie dla dostawców usług. Zaproponowane podejście umożliwia uwierzytelnianie użytkowników na podstawie nazwy konta i hasła oraz wymaga zainstalowania zaledwie jednego certyfikatu w urządzeniu końcowym.

Address

Piotr ZAWADZKI: Silesian University of Technology, Institute of Electronics,
ul. Akademicka 16, 44-100 Gliwice, Poland, Piotr.Zawadzki@polsl.pl .