International Journal of Electronics Engineering, 3 (1), 2011, pp. 99–102

Security Solutions in VOIP Applications

Anuj Kumar

Research Scholar, Singhania University, Pacheri Badi, Jhunjhunu, Rajasthan, INDIA,

E-mail: anuj.k.er@gmail.com

Abstract: Among the multimedia services over IP networks, voice/telephone communications are gaining greater and greater interest, either for personal and business applications. As a consequence, security solutions to guarantee privacy of communications and users' authentication become a necessary feature. This paper examines the state-of-the-art of security solutions for Voice over IP applications: the actual availability and support of such security facilities, through open source clients and servers, are discussed and tested experimentally. Moreover, the impact of security mechanisms on the perceived quality of the voice communications is evaluated, by means of suited software tools, that allow to simulate different traffic conditions and bandwidth availability, and to compare the corresponding audio Mean Opinion Score values obtained.

1. INTRODUCTION

Voice over IP (VoIP) applications allows the real time transport of voice/telephone communications over packet-switched networks based on the IP protocol stack. The same IP network usually has to accommodate general purpose data traffic and real time streams, thus having to face specific Quality of Service (QoS) requirements for the delivery of real time traffic. On the other hand, the convergence of voice and data networks brings a number of advantages, such as costs reduction and simplified network management. Besides QoS constraints, however, there are also other important issues to cope with, in a mixed data/voice traffic context: among them, users' authentication and privacy of their communications are of primary interest. VoIP applications may use a number of different protocols for transporting real time traffic; they may also resort to different protocols for call signaling and management. Among the most widespread solutions, we can cite the Session Initiation Protocol (SIP) [1], the H.323 protocol [2], and the Inter Asterisk Exchange v.2 (IAX2) [3] protocol. The last one is specifically devoted and optimized for VoIP applications, whereas former solutions are currently applied also to other real time communications over IP, like videoconferencing or video communications. The main components of a SIP architecture are well known; SIP is based on the use of textual messages in the clear, that are very similar to HTTP messages, aimed at session management and parameters negotiation among SIP clients (through the Session Description Protocol, SDP [4]). The H.323 protocol is a complex specification conceived for multimedia communications over packet switched networks; it has been reviewed during time, in order to become suited for VoIP applications. However, H.323 is currently the less adopted solution in the context of voice communications over IP networks, due to its complexity, and time-consuming connection setup phase. IAX2 is defined as a peer-to-peer protocol, meaning that users involved in a IAX2 communication maintain their own state machines associated to the protocol operations: IAX2, a specific VoIP oriented solution, manages both the signaling and transport of multimedia streams. Even if developed to be compatible with a number of different multimedia applications, this protocol is optimized for VoIP communications, thanks to reduced associated overhead and bandwidth requirements; in the IAX2 approach, signaling and data transfer are multiplexed in a single User Datagram Protocol (UDP) association between internet hosts. A specific feature of the IAX2 protocol, which is of interest for the aims of our evaluation, is the support of a symmetric ciphering scheme to provide confidentiality of the data frames exchanged between communicating peers. This feature is better discussed in the following.

In general, most of the SIP messages exchanged during a call-setup phase are not authenticated, and usually the communicating parties do not apply source control; on the other hand, H.323 can suffer from attacks performed by unauthenticated entities, and so called Man-in-the-Middle (MitM) attacks. It is clear that the actual impact of any security attack should be evaluated according with the communication scenario of interest; consequently, the security solutions aimed at counteract such attacks should be properly assessed, by a trade off among resource requirements and performance expected.

The paper is organized as follows: Section 2 briefly outlines available security mechanisms for VoIP systems; Section 3 discusses the security set up during signaling exchange, according with the different protocols adopted. Finally, Section 4 presents some experimental results about the impact of security operations on several parameters affecting the quality of a VoIP communication.

2. SECURITY MECHANISMS FOR VOIP SYSTEMS

Among the signaling and management protocols previously cited, IAX2 implements the Advanced Encryption Standard (AES), that is applied to 128-bit data blocks, and may use a secret key of length varying from 128 to 256 bits. IAX2 allows the application of the AES to any frame exchanged during a call, if such an option is properly signaled in a specific message, the NEW message, sent in the clear to start a call. Such an option is activated by adding a proper Information Element (IE) to the IAX2 message, that is the ENCRYPTION IE. In case the called peer is able to support data encryption, it replies with an AUTHREQ message sent in the clear, and carrying an ENCRYPTION IE; all the messages following this initial handshake can be ciphered. If the called peer is not able to support data encryption, its AUTHREQ message will not contain any ENCRYPTION IE; the caller can decide whether to close the session (by sending an HANGUP message) or to carry on with it, but without any encryption applied. The 128 bit shared key, used for data encryption, is computed from the CHALLENGE IE of the AUTHREQ message, which is concatenated to one of the shared passwords and hashed by means of the MD5 algorithm. The encryption engine is applied to Full and Mini frames, two different categories of frames used by the IAX2 protocol, with the exception of the Call Number fields carried in the header. More in details, IAX2 messages are usually classified as reliable or best effort messages. Reliable messages, the delivery of which is ensured by the protocol, are called Full frames: besides carrying the users' IDs, they may also show other attributes, the IEs. Best effort messages are called Mini frames or Meta frames: usually they are short messages carrying the user ID only, without any IEs. IAX2 frames may be subjected to padding, if needed, before AES can be applied. The format of IAX2 Full and Mini frames is shown in Fig. 1.

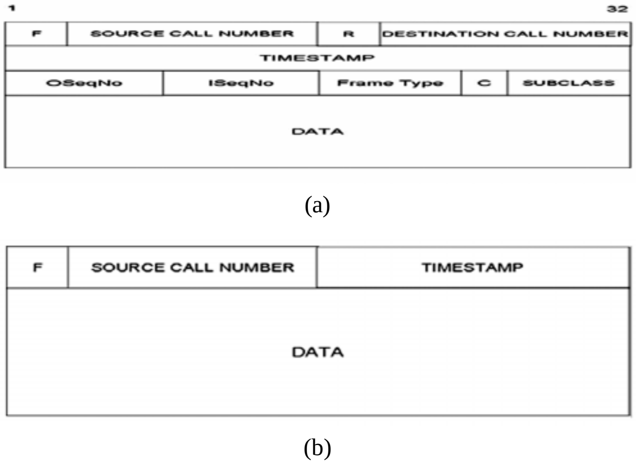
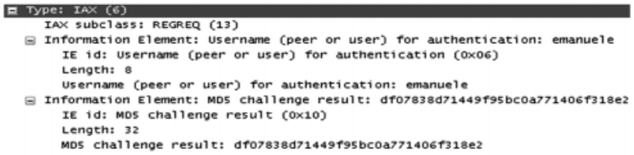


Fig. 1: IAX2 Message Format: (a) Full Frame, (b) Mini Frame

A suited security strategy, that may compensate possible weaknesses in the signaling and management protocol adopted in a real time communication over an IP network, relies on the Secure Real Time Protocol (SRTP), which provides symmetric encryption of the transferred user data. While both Real Time Protocol (RTP) and Real Time Control Protocol (RTCP) do not supply any protection against sniffing and data manipulation, SRTP supports media stream ciphering, caller authentication, data integrity check, and protection against replay attacks, by means of a key hierarchy (managed through Security Descriptions in SIP messages, or the MIKEY protocol [5]) and a Pseudo Random Function (PRF). The encryption operation is applied to the RTP data stream, whereas key exchange takes place during the signaling phase, that is made secure by application of SIPS, i.e. SIP over Secure Socket Layer (SSL)/Transport Layer Security (TLS). Despite the availability of possibly suited security countermeasures, ranging fromapplication specific solutions, to general purpose security architectures (e.g. VPNs, IPSec networks, and so on), most of the communications currently exchanged through VoIP connections are in the clear, and may be easily subjected to privacy violations. This holds also because most of the available software VoIP clients (the so called Soft phones) usually do not implement any security mechanisms, or, when present, they result to be quite complicated to be set up. In the framework of the proposed work, the state-of- the-art in confidentiality and authentication mechanisms that may be implemented in VoIP applications has been investigated, and an experimental evaluation of security countermeasures impact on the perceived quality of VoIP conversations developed. This point has been carried out through the adoption of proper software tools able to simulate different network congestion conditions, to get an objective indication of the corresponding Mean Opinion Score values (adopted as a quality metric), in presence of security schemes enabled or not. This last issue is quite unexplored in the available technical literature, and may be of great help in the design and provisioning of secure VoIP services.

3. SECURITY SET UP IN THE SIGNALING PROTOCOLS

Each signaling protocol has its own different configurations for the set up of security mechanisms. This paper deals with security configurations in the IAX2 and SIP protocols, the ones most adopted in VoIP applications. The effective set up of security options, in each case, has been verified through network traffic analyses, to confirm the real application of security mechanisms by the signaling protocol, as shown in Fig.2 for the IAX2 protocol.



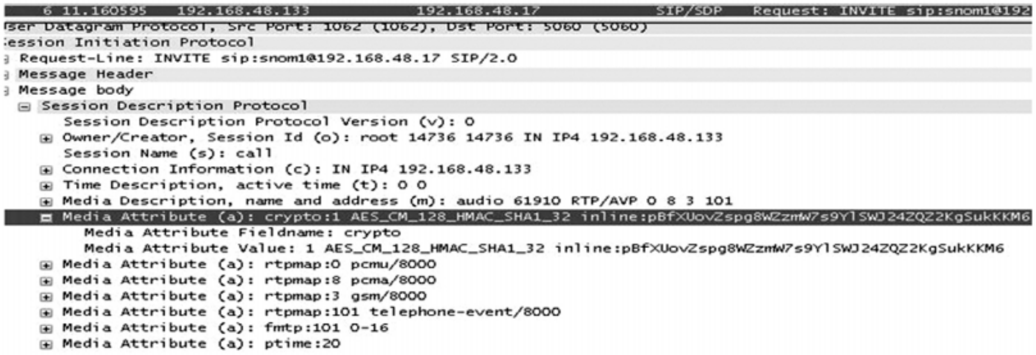
```
⊟ Type: IAX (6)
    IAX subclass: REGREQ (13)
  ⊟ Information Element: Username (peer or user) for authentication: emanuele
      IE id: Username (peer or user) for authentication (0x06)
      Length: 8
      Username (peer or user) for authentication: emanuele
  ⊟ Information Element: MD5 challenge result: df07838d71449f95bc0a771406f318e2
      IE id: MD5 challenge result (0x10)
      Length: 32
      MD5 challenge result: df07838d71449f95bc0a771406f318e2
```

Fig. 2: IAX2 Message Exchanged During the user Authentication Phase

In the case of security features developed for the SIP protocol, problems emerged during the corresponding configuration phase, when trying to make SIP working jointly with the Asterisk VoIP server adopted as the reference platform. Within Asterisk, SIP authentication is supported by using the Username/Password credentials associated to each client, that are to be properly set in the VoIP server inside the sip. Conf configuration file, as shown in the following code extract:

```
[Client Name]
…
Username = sjphone1
Secret = sjphone1
Host = dynamic
```

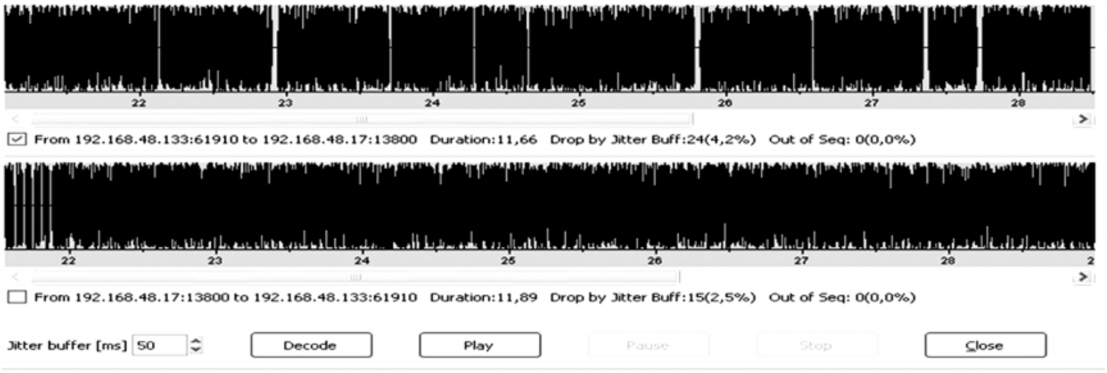
Several difficulties have been experienced also in finding available free soft phones that can effectively support the activation of security options during a SIP call set up and exchange. Among them, the SNOM360 SIP phone allows to exchange users' credentials through the Descriptions scheme, and supports the possibility of setting up an SRTP session. Fig. 3 shows a subset of the signaling exchange between a SNOM360 client and the Asterisk VoIP server, where the Crypto attribute is evidenced. The ciphering suite selected by the communicating entities is based on AES 128, with HMAC/SHA-1 as the hashing algorithm.



```
     6 11.160595    192.168.48.133             192.168.48.17             SIP/SDP   Request: INVITE sip:snom1@192
ser Datagram Protocol, Src Port: 1062 (1062), Dst Port: 5060 (5060)
ession Initiation Protocol
 Request-Line: INVITE sip:snom1@192.168.48.17 SIP/2.0
 Message Header
 Message body
  ⊟ Session Description Protocol
       Session Description Protocol Version (v): 0
    ⊞ Owner/Creator, Session Id (o): root 14736 14736 IN IP4 192.168.48.133
       Session Name (s): call
    ⊞ Connection Information (c): IN IP4 192.168.48.133
    ⊞ Time Description, active time (t): 0 0
    ⊞ Media Description, name and address (m): audio 61910 RTP/AVP 0 8 3 101
    ⊟ Media Attribute (a): crypto:1 AES_CM_128_HMAC_SHA1_32 inline:pBfXUovZspg8WZzmW7s9Y1SWJ24ZQZ2KgSukKKM6
         Media Attribute Fieldname: crypto
         Media Attribute Value: 1 AES_CM_128_HMAC_SHA1_32 inline:pBfXUovZspg8WZzmW7s9Y1SWJ24ZQZ2KgSukKKM6
    ⊞ Media Attribute (a): rtpmap:0 pcmu/8000
    ⊞ Media Attribute (a): rtpmap:8 pcma/8000
    ⊞ Media Attribute (a): rtpmap:3 gsm/8000
    ⊞ Media Attribute (a): rtpmap:101 telephone-event/8000
    ⊞ Media Attribute (a): fmtp:101 0-16
    ⊞ Media Attribute (a): ptime:20
```

Fig. 3: Crypto Attribute, SNOM360 Soft Phone

Experimental tests, performed by means of software traffic monitoring tools able to decode the captured audio samples, confirm that ciphering effectively protects audio data from sniffing (Fig. 4 shows captured audio samples during a ciphered connection), and that Asterisk cannot support bridging when selecting SIP as the signaling protocol, with a cipher algorithm used to protect the exchanged information.



☑ From 192.168.48.133:61910 to 192.168.48.17:13800 Duration:11,66 Drop by Jitter Buff:24(4,2%) Out of Seq: 0(0,0%)

□ From 192.168.48.17:13800 to 192.168.48.133:61910 Duration:11,89 Drop by Jitter Buff:15(2,5%) Out of Seq: 0(0,0%)

Fig. 4: Audio Captured from a Ciphered SIP VoIP Session

Preliminary tests on the MIKEY protocol, for the real time management of cryptographic keys in SRTP, showed that MIKEY can support a pre-shared key mode, a public key mode, and a Diffie-Hellman mode for a secure symmetric key exchange. However, when testing the soft phone implementing MIKEY on the Win32 O.S., several crashes have been evidenced and the application cannot be used effectively.

3. IMPACT OF SECURITY ON VOIP COMMUNICATIONS QUALITY

Several network conditions, due to different values of delay time, average jitter, and packet losses, have been simulated in order to evaluate quality variations of the VoIP communications, due to the activation, or not, of the security options available. Resulting objective quality is expressed through the MOS parameter that can vary from 1.0 (not recommended) to 4.4 (very satisfactory); a reference MOS value representing an acceptable quality equals 3.5 (Ref value, in the following graphs). Fig. 5 shows a first example of the experimental results obtained: the measured MOS values (in case of a GSM6.1 audio codec) are affected not only by the percent amount of packet losses in the network, but also by the application, or not, of the SRTP protocol to ensure confidentiality of the exchanged audio data. As reported, when the SRTP option is activated, a lower amount of packet losses can be tolerated on the network, at a parity of the MOS requested by the VoIP communication. If we consider the joint effect of the delay time and the SRTP operations, the results shown in Fig. 6 are obtained, on average. The maximum delay time considered for testing purposes amounts to 60 ms. With respect to the non ciphered case, when SRTP is used, the MOS values obtained get smaller, and the quality provided is acceptable if the delay time is lower than 60 ms, for a bandwidth availability of at least 64 kbps.

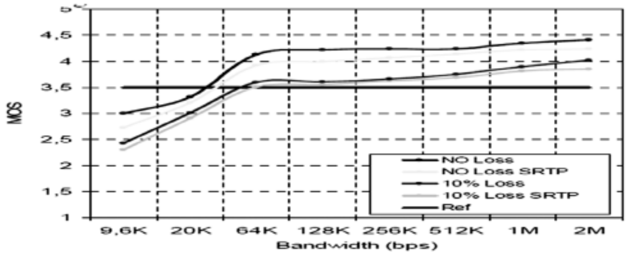


Fig. 5: MOS Variations Due to Losses and Ciphering (GSM6.1)

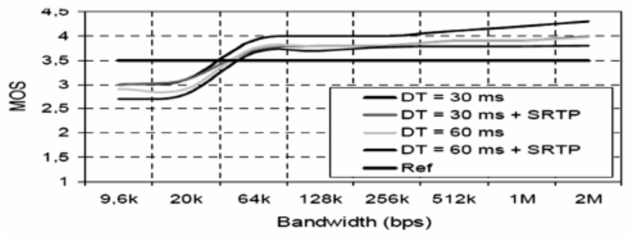


Fig. 6: MOS Variations Due to Delay Time and Ciphering (GSM6.1)

Jitter and other disturbances have been considered under the definition of a corruption rate, that can vary between 5% and 10% of the total amount of packets exchanged, meaning that 5 or 10% of all the packets were somehow corrupted. Fig. 7 shows the effects of corruption on the communication quality, when SRTP is applied or not. Again, in presence of the SRTP functionalities, the MOS values obtained are smaller, at a parity of the network conditions. Similar results can be obtained when testing a G.711 audio codec, but with reduced performances, due to the lower compression rate obtainable. Table I reports the final comparison between GSM6.1 and G.711 codec's, tested under similar conditions.

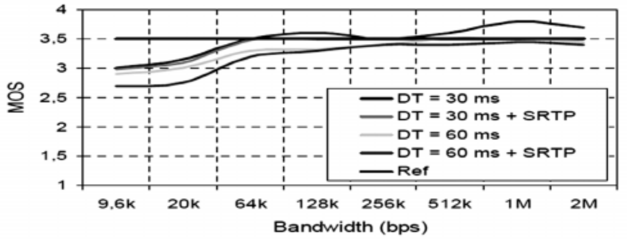


Fig. 7: MOS Variations Due to Corruption and Ciphering (GSM6.1)

Table 1
GSM6.1 and G.711 Performance Comparison

	GSM6.1		*G.7.11*	
	Plain	*SRTP*	*Plain*	*SRTP*
Min Bandwidth	30 Kbps	34 kbps	84 kbps	88 kbps
Max Packet Loss	10%-15%	5%-10%	15%-20%	10%
Max Delay Time	120 ms	90 ms	120-150 ms	90-120 ms
Max Corruption	5%-10	5%	10%	10%

REFERENCES

[1] J. Rosenberg, H. Schulzrinne, et al., “SIP: Session Initiation Protocol", RFC 3261, June 2002, Available at http://www.ietf.org/rfc.html

[2] ITU-T Telecommunication Standardization Sector of ITU, Series H, “H.323-Packet-Based Multimedia Communications Systems", June 2006

[3] http://www.cornfed.com/iax.pdf

[4] F. Andreasen, M.Baugher, D. Wing, “Session Description Protocol (SDP)”, RFC4568, July 2006, Available at http://www.ietf.org/rfc.html

[5] J. Arkko, E. Carrara, F. Lindholm, M. Naslund, K. Norrman, “MIKEY: Multimedia Internet Keying”, RFC 3830, Sep. 2004, Available at http://www.ietf.org/rfc.html