SETIT 2005
3rd International Conference: **S**ciences of **E**lectronic, **T**echnologies of **I**nformation and **T**elecommunications
March 27-31, 2005 – TUNISIA

Critical voice network security analysis and new approach for securing Voice over IP Communications

Carole Bassil*
Nicolas Rouhana**
& Ahmed Serhrouchni*

* *Computer and Networking Department, ENST Paris, CNRS*
** *Cimti, Faculty of engineering, Saint-Joseph University, Beirut, Lebanon*

carole.bassil@enst.fr
ahmed@enst.fr
nicolas.rouhana@fi.usj.edu.lb

Abstract: Voice networks evolved from the fixed traditional telephone system, to mobile and wireless networks and now towards a converged voice and data infrastructure. This convergence is based on the spread of the Internet Protocol, where VoIP is developing. Due to IP network characteristics, hackers are able to compromise and take control of different aspects of IP telephony such as signaling information and media packets. Security and privacy become mandatory requirements for this application area. IP telephony requires security services such as confidentiality, integrity, authentication, non-replay and non-repudiation. The available solutions are generic and do not respect voice specificities and constraints. Thus, QoS of the voice is affected by delay, jitter, and packet loss. In this paper, we present the security mechanisms as provided in the existing fixed, wireless and converged voice networks, outline major security weaknesses in these different environments and end this paper by a proposition to secure voice over IP packets drawing inspiration from the existing voice security solutions.

Keywords: Security mechanisms; GSM; Bluetooth; IPsec; SRTP, VoIP

1- Introduction

Plain Old Telephone System (POTS) is the oldest telephone system offering analog voice services. It was commonly considered as a secure network. However, phone tapping may simply involve the installation of a low cost capacitor and the snipping of a wire. Integrated Services Digital Network (ISDN) brings us closer to the goal of a ubiquitous multi-service network, integrating voice, data, and video services in a digital format over a common global network. It was subject to many standards, but no security standards were defined and implemented to protect the network from eavesdropping and accessing critical information carried over its channels. However, security related discussions have been reported by the Integrated OSI, ISDN and Security Program of the Computer Systems Laboratory at the National Institute of Standards and Technology [29]. Digital wireless voice transmission technologies were introduced by mobile networks such as the GSM [3] and Bluetooth [12]. Efforts were made at the IEEE 802.11 [17] working group to implement voice over Wi-Fi by defining new RFCs to assure security and QoS. Wireless technology, by its nature, violates fundamental security principles; it presents security caveats against attacks launched over the radio path. Currently, Internet took the leadership by providing value added services mainly IP telephony services at very low cost. Nevertheless, transferring critical information over communication infrastructure accessible to the public presents security vulnerabilities.

Various common security requirements [1] have to be met to secure a transmission:

- Authentication is the property of knowing that the claimed sender is in fact the actual sender.

- Privacy and Confidentiality are the properties of communicating such that the intended recipients know what was being sent but unintended parties cannot determine what was sent. These features are accomplished by ciphering the exchanged information.
- Integrity is the property of ensuring that data is transmitted from source to destination without alteration. This feature is fulfilled by a unidirectional hashing function.
- Non-repudiation is the property of a receiver being able to prove that the sender of some data did in fact send the data even though the sender might later desire to deny ever having sent that data.
- Non-replay is a protection against replay attacks which consist of memorizing the data by a third party then re-injecting it onto the network.
- Resource availability could also be considered as a security mechanism in order to minimize delays and service reject.

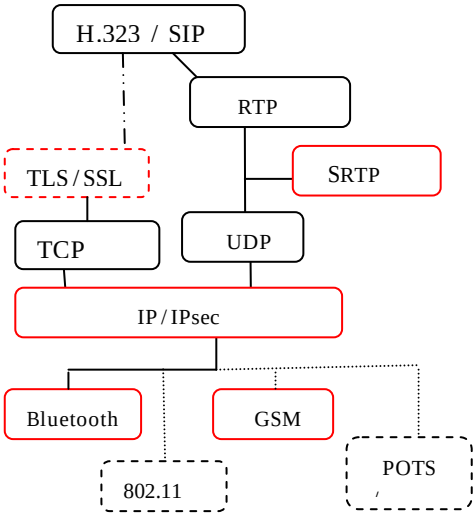

Figure 1 voice systems and security bricks

In this paper, we present different technologies that provide voice capabilities and security services. The goal of our investigation is to analyze these different voice platforms with regards to their security applications and understand whether these multiple algorithms and security protocols apply to the voice based on the withheld options and their performance impact on the voice quality. The results of the conducted analysis will lead us to propose a secure voice over IP solution inspired from the points of strengths of the existing deployed security services. Figure 1 shows different voice systems and security bricks. Thus, GSM [3] secures the voice communication over the radio link as outlined in section 2. Bluetooth [12] provides voice and security mechanisms as presented in section 3. Section 4 introduces the IPsec security framework standards specifically designed to operate at the network level, thus providing security to any application that runs over IP. The Real-time Transport Protocol [30] provides end-to-end network transport functions suitable for audio applications. The secure RTP profile, SRTP [16], provides security mechanisms to protect the RTP and its control traffic as described in section 5. Voice capabilities in IP-based networks are offered through H.323 [6] and SIP [20] standards. The related security requirement as defined in the H.235 [7] security recommendation of H.323 and SIP security will be introduced in section 6. In Section 7, we introduce a solution to secure voice packets carried over IP by combining the confidentiality, integrity and non-replay mechanisms provided by SRTP along with our proposition to provide authentication and non-repudiation mechanisms. Section 8 concludes the paper and summarizes our findings.

2- Security in GSM

Global System for Mobile Communications (GSM) [3] is the most popular mobile telephone network. Using the radio path for communicating with the user makes the GSM network sensitive to:

- Misuse of its resources by unauthorized persons using manipulated mobile stations.
- Eavesdropping on the information which is exchanged on the radio path

In order to protect the system, different security features were reinforced mainly authentication and confidentiality as shown in Figure 2. *Authentication [2]* is used to identify the user to the network operator, based on challenge-response encrypted techniques. A specific authentication algorithm (A3) calculates a signature that it sends to the network based on the individual key (K_i) registered on the user SIM card and the challenge sent by the network. The network compares it with the one provided by its database to authenticate the user. *Confidentiality [2]* is provided by the challenge sent by the network, the user's Ki and the specific A8 algorithm to generate a session key (K_c). Each frame crossing the radio link will be encrypted with a different key generated by the A5 encryption algorithm. This algorithm is initialized with the session key (K_c) and the number of frame to be encrypted generating a different key for each frame. *Integrity* is not provided within GSM based on cryptographic algorithm. *Non-repudiation* and *non-replay* are not provided by GSM. However, billing invoices could be used as a proof against repudiation.

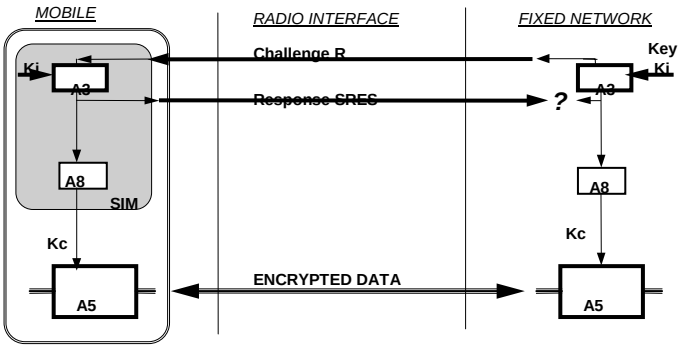

Figure 2 Authentication and Encryption in GSM

Note: For security reasons, A3, A8 and A5/1 and A5/2 algorithms were not available to the public [3]

GSM security weaknesses:

- In a GSM network, the functions described above are applied only on the radio link between the mobile station and the network. Communications and signaling within the core network and connection with the fixed network are not protected.

- The implemented comp128 authentication algorithm (A3/A8 algorithms) was broken [22].
- The encryption A5 algorithm showed vulnerabilities with attacks [22] conducted against its implementation, especially the A5/2 [24].

3- Security in Bluetooth

Bluetooth is a wireless technology that realizes small personal networks. This technology implements a synchronized time technique where time is divided into slots. Security measures are defined at the link layer within Bluetooth. This link layer security is based on different algorithms as shown in Figure 3 and explained below.

Authentication [4] is based on challenge-response principle using a specific E1 algorithm. The receiver's E1 algorithm uses the challenge sent by the correspondent, its Bluetooth address and the current link key to generate a digest to send to the transmitter. The verifier will execute the same operation to confirm the answer received. The above process will be repeated in both directions to authenticate both parties.

Confidentiality [4] is provided by encrypting the payload based on the E0 algorithm. A unique encryption key is generated for each session and from which keys per packet are derived. This encryption key is calculated and generated as the digest of the combination of a random number, the link key and a product of the authentication procedure using the E3 algorithm. In Bluetooth all data carried on its different links (except baseband headers) are encrypted with the E0 when encryption applies. *Integrity*, *Non-replay* and *Non-repudiation* are not provided by Bluetooth.

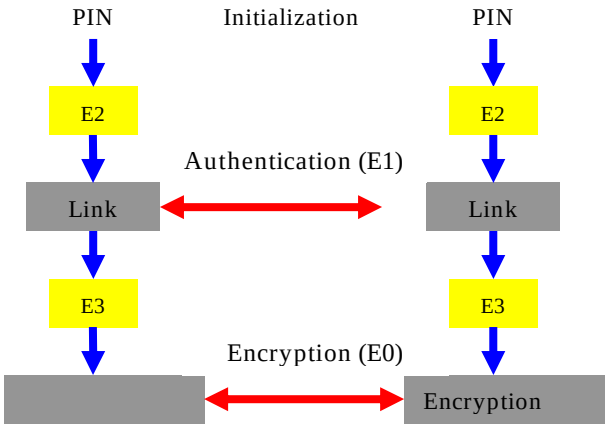

Figure 3: Link security Architecture in Bluetooth

Bluetooth security weaknesses:

- Only the equipment is authenticated.
- The different keys could be stolen leading to eavesdropping or even impersonating devices.
- Bluetooth security is provided basically on the access level, what would be the impact on the voice quality and performance of securing an end-to-end connection (Bluetooth, GSM, upper layers)?

4 – Security in IPsec

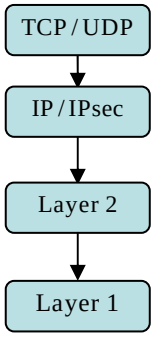

Figure 4 IPsec in the stack

IPsec is a suite of protocols designed to secure communication at the network level. These protocols are in constant evolution since 1995. New drafts are proposed within the IETF working groups. IPsec proposes two security protocols: **A**uthentication **H**eader [13] and **E**ncapsulating **S**ecurity **P**ayload [14].

Figure 4 shows the position of IPsec in the protocol stack. The security mechanisms deployed in AH and ESP are detailed as follows:

Authentication is provided by the AH header mechanisms using a hash function with a secret key. The algorithm used is negotiated within a security association established between sender and receiver. The default authentication algorithm is HMAC-MD5 [35, 36]. Using AH will increase the processing time and the delay [13]. Authentication provided by ESP is very similar to AH. The only difference between AH and ESP is in the ESP Transport mode where the IP header is sent in clear text [14].

Confidentiality is provided by encrypting the payload of an ESP packet. The cryptography algorithm used is negotiated within a security association. ESP is designed to work with a symmetric cryptography algorithm (DES-CBC [37] by default, AES [21] recently). Each IP packet should carry specific data for cryptography synchronization used for decryption [14]. ESP implementation introduces more complexity at the user level. AH does not provide encryption. *Integrity* in connectionless mode is obtained by calculating an Integrity Check Value. The algorithm used for calculating the ICV is based on the symmetric key algorithm (DES) [23] or on a one-way hash function (MD5 [36] or SHA1 [38]). The same process is used with ESP to perform the integrity check of the conveyed data. *Non-replay* is provided by the sequence number included in the AH and ESP headers. This number is incremented at each transmission of an IPsec packet. This function is negotiated at the receiver request in a security association. *Non-repudiation* is present by using the RSA asymmetric algorithm for authentication with AH [13], the transmitter and receiver keys are used to calculate the authentication data. This feature is not available when applying ESP security mechanism only.

Security Association negotiates security parameters such as authentication and encryption algorithms, keys, initialization vectors, counters, etc. between two entities.

IPsec security weaknesses

Two main factors affect voice traffic when IPsec is used:

- The increased packet size is due to headers added to the original IP packet
- The time required to encrypt payload and headers and the construction of new ones is high

- Authentication covers machines only (logical address). Users are not identified.
- IPsec crypto-engine is a serious bottleneck in the transmission of real-time traffic. It is impossible to control packet access to the crypto-engine [15].

5. Security in SRTP

SRTP (Secure Real-time Transport Protocol) [16] is a security of RTP standard which provides confidentiality, integrity, and authentication and replay protection. SRTP encrypts the payload while leaving the packet header as clear text. SRTP is independent from the underlying layers used by RTP. SRTP is characterized by a high throughput leading to minimize the processing time and low packet expansion. The following is the security services provided by SRTP.

Authentication is based on a hash function with a key invoked to authenticate the header and payload of the RTP packets. A Message Authentication Code (MAC) is appended to the end of the packet and verified by the receiver by re-computing this MAC using the same process. The default algorithm used to provide authentication and integrity is the HMAC-SHA1 [35, 38]. The source messages authentication is provided in a peer-to-peer communication only.

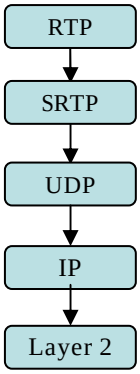

Figure 5 SRTP in the stack

Confidentiality: the payload and the authentication tag appended to the RTP packet are encrypted by the transmitter and the receiver using the same session key. One master key is sufficient to insure the confidentiality and integrity of the RTP/RTCP packets. This is performed using a derivation key function which creates session keys derived from the master key. SRTP uses a key-stream approach to encipher the flow of information. Key-stream generation is accomplished by the AES [31] encryption algorithm. *Integrity* is obtained using the hash function provided with the authentication mechanism. *Replay protection* is assured if integrity is enabled. Each SRTP receiver maintains a replay list which indicates the indices of all authenticated received packets using a sliding window technique. After authenticating a packet, this list is updated with the new indices. *Non-repudiation* is not developed within SRTP.

SRTP security weaknesses:

- We need for a separate management key (e.g. IKE, ISAKMP/Oakley, Kerberos or point-to-point mechanisms such as Diffie-Hellman algorithm).
- We need to modify the protocol in all existing IP phones
- There is no user authentication in unicast, multicast and RTP group sessions.
- Only the source origin of the data is authenticated.
- The RTP headers are sent in clear text to allow header compression, which leaves certain fields available to attacks.

6- Security in VoIP

VoIP can be defined as the ability to make telephone calls over IP-based data networks with a suitable quality of service (QoS) and a much superior cost/benefit. The VoIP standards used are H.323 and SIP.

Figure 3 shows the different security protocols used to carry and secure VoIP network. H.235 [7] defines the security requirement in H.323 [6] environment. TLS [5] is used to secure the signaling channel. SIP [26] may use TLS/SSL [5, 19]. SRTP can be used to secure voice packets. IPsec can be used as an alternative. In the following section, we are going to outline the security mechanisms provided by H.323 and SIP voice over IP standards.

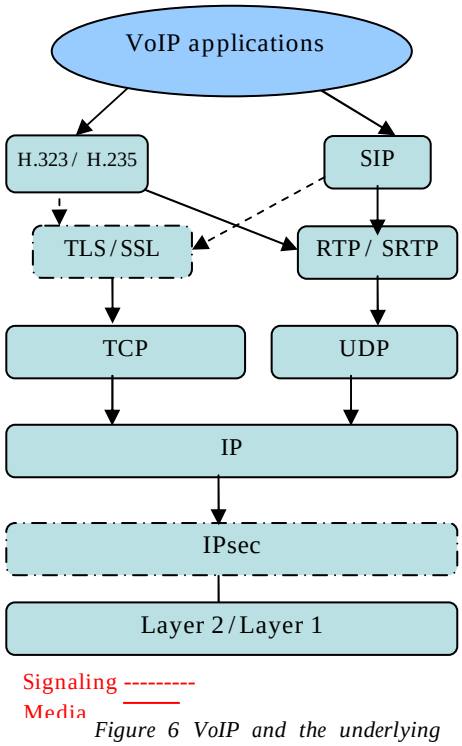

Figure 6 VoIP and the underlying security bricks

6.1 Security in SIP

The Session Initiation Protocol (SIP) [20] is a text-based, application layer control (signaling) protocol for creating, modifying and terminating multimedia and voice sessions with one or more participants. The fundamental network-security services required for SIP are provided in [26].

Authentication is provided hop-by-hop by TLS or IPsec. Also, SIP can make use of any transport-layer or HTTP-like security mechanism. SIP also defines end-to-end authentication using either PGP [27] or S/MIME [28]. *Confidentiality* is supported with two forms of encryption. The first method is end-to-end encryption of certain sensitive header fields and the message body using either PGP or S/MIME. Hop-by-hop encryption is the second method that can be used to hide the route a message takes, as well as preventing eavesdropping. SIP could use any transport-layer or HTTP-like security mechanism for encryption. Keys for media encryption

are conveyed using SDP [25]. Transport- or network-layer security encrypts SIP signaling traffic, guaranteeing message confidentiality. *Integrity* is not supported by SIP. However, transport- or network-layer security could be used to guarantee SIP signaling message integrity. *Replay protection* is provided with specific SIP headers [20] to identify and order transactions. *Non-repudiation* could be provided with IPsec.

6.2 Security in H.323

The H.323 standard is made of important building blocks constituting a foundation for audio, video and data communications across IP-based networks. The signaling information is transported reliably over TCP. Voice information is carried by RTP/RTCP packets over UDP. H.235 [7] is part of H.323 building blocks that define security for VoIP as follows:

Authentication [7] is based on two types of concepts: symmetric encryption-based or shared secret known as "subscription" with three variations:

- password-based with symmetric encryption;
- password-based with hashing (also symmetric);
- certificate-based with signatures (asymmetric) to authenticate the use itself.

As a third option, authentication may be accomplished by TLS [5] or IPSEC [13]. Authentication is brought out during call establishment connection on the signaling channel. The information carried on the signaling channel can secured with TLS. *Confidentiality* can be provided for call control and media channels [7] to protect the data carried on these logical channels. The required cryptography key used for encrypting media channels can be carried on a specific H.235 logical channel. *Integrity* of the exchanged information over all logical channels is provided by using hash functions in conjunction with the deployed authentication mechanisms [8], [10] and [11]. *Non-Replay* was not considered and developed in H.235 recommendation. However, a sequence number and timestamp fields of packet headers could be used to provide protection against replay attacks. *Non-repudiation [9]* could be provided using a digital signature in conjunction with a one-way hash function such as MD5 or SHA1.

Voice over IP security weaknesses:

- There is a lack of security in VoIP implementation by the IP phones vendors.
- Most of firewalls are not VoIP aware.
- The signaling/control and the media data might be the major target of attacks.
- Some of the proposed security solution is not suitable for voice such as SSL [19] and TLS which secures TCP traffic while voice packets are carried by UDP packets.

7- Proposition for securing Voice over IP

Every security solution described earlier possesses advantages and inconveniences regarding its deployment to secure voice transmission. Our purpose is to analyze the pros and cons, take inspiration from the strong points of each solution and propose a solution to secure voice packets over IP-based networks.

7.1 Security analysis:

POTS/ISDN services provide line number identification (Caller Line ID) through signaling messages. Bluetooth allows mutual authentication of communicating devices based on their Bluetooth address thus providing only device authentication. GSM uses a unique user identifier attributed by the network operator and registered on the SIM card to identify and authenticate the user based on challenge-response operations. Data origin authentication is offered by SRTP and IPsec authentication mechanisms while user identity is *authenticated* by means of digital certificates or digital signature as suggested by VoIP standards. However, digital certificates are heavy to implement with voice since it relies on public key cryptography which requires the exponentiation of large numbers, a computationally intensive process which limits their speed. For those reasons applying public key algorithm to secure voice transmission will introduce delays and affects the overall performance. Therefore, a new authentication and identification mechanisms should be proposed which respects the constraints and quality requirement of the voice nature.

Confidentiality based on different encryption techniques was used to encrypt a point-to-point communication over traditional fixed voice networks. GSM mobile network introduced a new technique to insure the privacy of the communication over the radio link where each transmitted packet is encrypted with a different key calculated by a specific encryption algorithm to protect the data from attacks. Bluetooth is using a specific stream cipher algorithm to secure the wireless path taking into account devices resources and network characteristics. Amongst the encryption algorithms deployed in the IP based networks, Advanced Encryption Standard (AES) is the most simple, flexible, efficient (computational efficiency and less memory requirements on software and hardware, including smart cards) and secure (key length 128, 192 and 256 bits) symmetric algorithm.

Integrity is provided by a hash function calculated over the original message or a hashed digital signature performed using public key algorithm. *Replay list* with a sliding window approach based on the sequence numbers and timestamp as defined by SRTP is a good technique to avoid replay attacks.

Non-repudiation functionality is only provided by public key algorithm while signing the messages with a digital signature along with a timestamp. Since Public key algorithms are very heavy to implement with voice transmission, different approach should be considered to avoid deny of reception or deny of participation in a conference call.

Performance issues: Various factors influence signal delay during a VoIP transmission.

- The time spent by the CODEC may vary between 0.75 – 30 ms, depending on the coding schemes adopted and the quality of the reproduced signal.
- Segmentation and encapsulation within IP packets.
- Queuing delay (i.e., time spend by a packet in the router buffers before routing) may add up to 30 ms.
- Jitter delay in the range of 40 – 70 ms is introduced by buffering the arriving packets to be delivered at a

uniform rate. Buffering is necessary to eliminate the variation of the delivery rate caused mostly by queuing time due to network load.

Because of such timing constraints, voice packets are small (10-50 bytes payload length [18]) in order to guarantee that all the above mentioned operations can be performed within the given time constraint. The payload length would be a compromised value between the bandwidth and the voice quality.

Performance of securing VoIP with IPsec: the below table shows an increase in the size of the packets when IPsec is used to secure VoIP transmission, reaching up to 130 bytes of total length depending on the cryptographic services requested. This leads to a decrease in the number of effective phone calls [18] by half for a 128kbps line with 50 packets per second (pps). When using IPsec, the payload to the total packet length ratio is decreased by half from 34% to 18%, as shown in table 1. The increase in the packet size has negative effects delays and on bandwidth usage. Such a result is not surprising since the time required by encryption is not negligible.

Implementing IPsec to secure VoIP traffic adds delays and overheads which will decrease the performance of the voice and affects the QoS of the call.

Packet Type	Header (Bytes)	Packet length (bytes)	Ratio	Number of calls
IP	40	60	34%	4.7
IPsec DES	82	102	20%	3.2
IPsec 3DES + SHA	94	114	18%	2.6

Table 1 - header and total packet length, payload to total packet length ratio, number of calls on a 128 kbps link at 50 pps using G.729 codec

Performance when securing VoIP with SRTP:
SRTP adds 4 bytes to the RTP packet length to encrypt the payload of the RTP carrying voice information. If we calculate the possible number of calls [18] on a 128 kbps link at 50pps using SRTP, we should obtain the following results:

Packets Type	Header (bytes)	Packet length (bytes)	Ratio	Number of calls
RTP	40	60	34%	4.7
SRTP with AES	44	64	31%	4.5

Table 2- header and total packet length, payload to total packet length ratio, number of calls on a 128 kbps link at 50 pps using G.729 codec

We can notice that SRTP does not affect the total number of effective calls. The payload to total packet ratio does not indicate a major decrease which means that the overhead introduced by SRTP is very small compared to the overhead introduced by IPsec (table 1). It is possible with SRTP to compress the IP/UDP/RTP headers with specific compression algorithm. Compression will help minimizing delays during transmission and optimizing the performance of the applied security mechanisms.

Our proposed solution as highlighted in table 3 will be based on the confidentiality, integrity and replay protection security mechanisms deployed by SRTP. The GSM authentication mechanism will be reconsidered to operate within IP network; non-repudiation will be inspired from that defined within [9].

	GSM	BT	IPsec	VoIP	SRTP
Authentication	+	+	+	+	+
Integrity	-	-	+	+	+
Confidentiality	+	+	+	+	+
Non-Repudiation	-	-	-	+	-
Replay protection	-	-	-	-	+

Table 3 –security mechanisms used in our proposed solution. (+: service provided, -: service not available)

7.2 Proposed solution:

In order to provide a secure voice call between two communicating entities, it is necessary to take into consideration which type of call model and the related vulnerabilities of the infrastructures being used to carry out this voice communication.

Three different call models can be used in telephony. The first model is a communication established between two end users using the operator networks. We can give as an example two users connected between them through the traditional telephone or mobile network, or connected between them through the Internet for a voice over IP session or any combination of these networks. The second model can be a communication established between two remote users while passing through one intermediary equipment to which one of the two communicating users is connected. The remote user is connected directly to the operator network. This intermediary equipment can be the PABX of an enterprise, a voice over IP gateway or any other trusted entity as well. The third model is a communication established between two users each of one can be a subscriber of either a conventional PABX of an enterprise as well as a voice over IP gateway or any combination of two intermediary equipments.

As a call can pass through heterogeneous infrastructures, it is then necessary to have an abstraction of the underlying infrastructure and insure the security of the communication independently from the deployed techniques to establish the call.

A third trusted party can be necessary to realize a secure transmission. It can be responsible of the distribution of a secret key shared between the two actors while preserving this key from aggression. This trusted third party can play the role of arbitrator between two actors when disputes arose concerning the authenticity and the non-repudiation of the communication.

We propose to define the authentication mechanism with an associated non-repudiation feature to be deployed within the second telephony context model as described

above. We consider that the user directly connected to an intermediary equipment needs to be authenticated while the remote end user is a trusted entity which do not need authentication. Therefore the authentication mechanism as described within this paper is a one-way authentication. The authentication mechanism will be based on smart cards to authenticate the user. A trusted third party (Trusted Authentication Authority) manages the delivery of smard cards. Adding smart cards provides reinforced security and resistance to attacks. The other two models will be handled in further works.

Smart Cards: the usual approach is to store in a permanent memory of the machine the required information such as the subscription identifier in the shape of digital signature for instance. This smart card is a kind of key. If removed from the terminal, the latter cannot be used except for emergency calls, and that is to say it cannot be used for any service which will impact the subscriber's bill. These smart cards can hold besides the user subscription identifier user profiles in terms of abbreviated dialing numbers list with the correspondent alphanumeric index, etc. It can be protected by a password (typically 4-digit) similar to the PINs of credit cards.

Trusted Authentication Authority (TAA): Its role is to memorize and attribute for each VoIP subscriber a secret key (digital signature), delivered in the form of smart cards and to manage tokens to be used in a later stage for billing purposes. The Trusted Authentication Authority will handle all user information (name, address, phone number, a signature digitally signed by a public key algorithm) in its User Information Database (UID). A TAA may be associated with a Gatekeeper or implemented as a standalone module at the Internet Service Provider location. Multiple TAAs might be spread over the VoIP network managed by a Global Trusted Authentication Authority (GTAA). Since the TAA is considered as a trusted entity, it possesses a PKI certificates. It might use it, if necessary, to protect the token delivered to the remote user (Bob) from man-in-the-middle attacks.

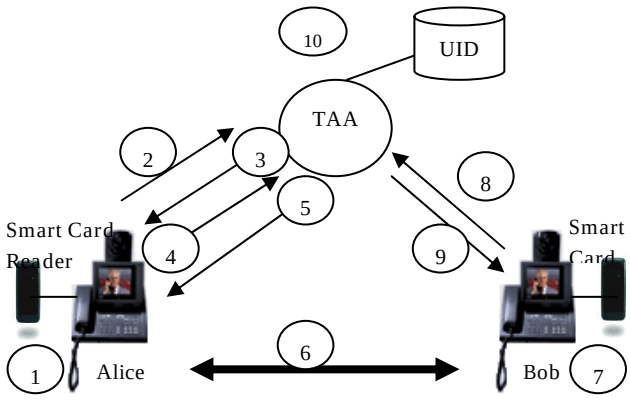

Fig. 7 Authentication mechanism based on smart

Authentication mechanism:
Figure 7 shows the proposed security mechanisms as explained below while figure 8 presents the exchanged messages during the authentication and non-repudiation phases.

1- Alice would like to initiate a phone call using an IP phone. She must first insert her authentication card in the smart card reader connected to the IP phone to be able to initiate phone calls through the IP network.
2- Upon inserting the smart card, an authentication request message will be initiated to the TAA asking for access grant.
3- The TAA sends Alice a random number (RAND) (figure 8).
4- Alice's smart card calculates a response based on the random number and her signature using a specific algorithm similar to that used in GSM SIM card. The resulting value will be hashed with a hash function running on the card, and then sends the resulting digest (HRES) (figure 8) to the TAA to be verified.
5- TAA will calculate the same digest based on the random number sent, the digital signature of the user stored in its database and the hash function. TAA will verify the digest received from the user with the one calculated. After verification of the result, the TAA will grant access to the user by sending a confirmation message along with a Token. The call will be initiated towards Bob's IP phone. The TAA considers Bob as an entity which does not to be authenticated; therefore it will deliver the Token to Bob. TAA will keep a record of the token delivered to both parties along with some parameters (e.g. starting time of the session, etc.) necessary for issuing billing invoices and for non-repudiation functionality.
6- Upon authenticating Alice, the voice call will be established between the two parties. Integrity and confidentiality features will be deployed to secure the voice communication transported by RTP packet according to the integrity and confidentiality provided by the SRTP standard.
7- A replay list should be managed at the receiver side to protect from replay attacks
8- During the exchange of signaling messages terminating the voice session between participants, the TAA must be informed of the session ending by returning the Token by both parties to keep records for future use.
9- TAA will acknowledge the reception of the Token to both parties.
10- TAA will keep records in a specific database related to the established session between the two parties Alice and Bob, in order to issue billing invoices and to avoid deny of the participant in the session insuring by that the completion of the non-repudiation functionality.

Note1: The delivered Token could be used by SRTP key management to generate session keys (master keys and slating keys, etc.). Integrating the token with SRTP key management is outside the scope of the current paper.

Note2: Authentication procedure is carried out over signaling messages during the establishment phase of the call. Therefore, the underlying signaling channel should be opened in a secure manner (i.e, using a well known TLS port).

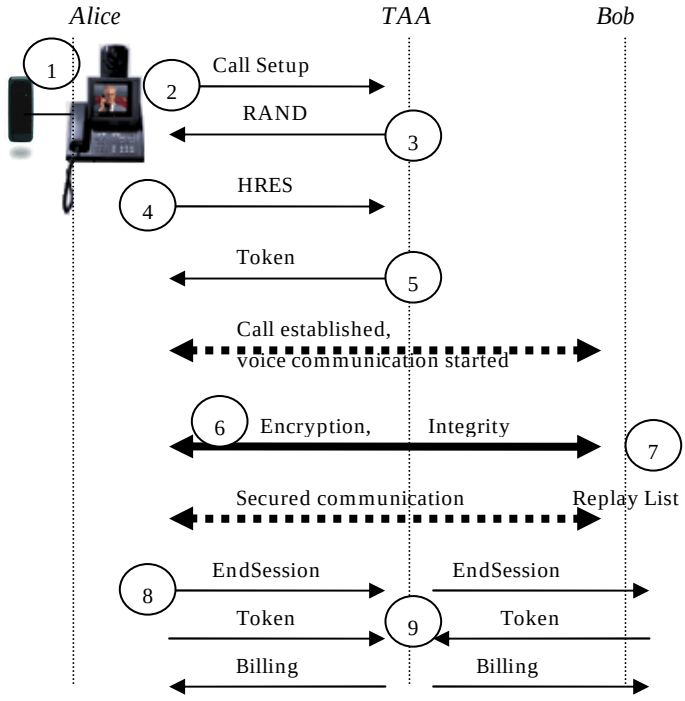

Figure 8 Authentication message exchange

Confidentiality: SRTP implements the strong AES encryption algorithm in two different variants [16]: AES in counter mode [32] or AES-f8 mode [33, 34]. The encryption transform maps the SRTP packet index and secret key into a pseudo-random key stream segment. Each key stream segment encrypts a single SRTP packet. The process of encryption would be as follows: generating the key stream segment corresponding to the packet, then bitwise exclusive-oring that key stream segment onto the payload of the RTP packet to produce the encryption Portion of the SRTP packets.

Integrity is deployed to protect packet alteration by untruthful parties. This integrity use a one-way hash function calculated over the entire SRTP packets resulting in a digest (authentication tag). This digest will be appended to the end of the packets. Integrity is provided to point-to-point communication as well as multicast communications and group sessions. RTP headers include sequence numbers which can be used to provide replay protection. Secure replay protection is only possible when integrity protection is present.

Non-replay: Replay attacks are avoided by a replay list, maintained by the receiver only containing indices of recently received and authenticated SRTP packets as follows: each SRTP receiver maintains a Replay List, which conceptually contains the indices of all of the packets which have been received and authenticated. In practice, the list can use a "sliding window" approach, so that a fixed amount of storage suffices for replay protection. The receiver checks the index of an incoming packet against the replay list and the window. Only packets with index ahead of the window, or, inside the window but not already received, shall be accepted. After the packet has been authenticated, (if necessary the window is first moved ahead), the replay list shall be updated with the new index. The Replay List can be efficiently implemented by using a bitmap to represent which packets have been received, as described in the Security Architecture for IP [1].

Non-repudiation:
Since the TAA is managing the authentication sessions and the delivery of token to parties registered in its database and participating in a particular secured VoIP communication, it will keep records of the digest received and the delivered token related to each participant in the voice session for future use to protect from the denial of participation of any involved participant in a specific voice call. At the end of the session, each user will send back the token to the TAA. The TAA will be able to issue billing invoices based on each user's profile. Thus, non-repudiation will be reinforced by the records kept and the billing invoices issued for each recipient.

The main challenge with smart cards is its integration within the actual IP phones. Specific efforts should be deployed to use smart cards with the existing VoIP standards and to interoperate with the deployed Trusted Authentication Authority to handle all the signaling information used during this authentication and non-repudiation scheme (i.e. open a secured transport channel, billing, etc.).

8- Conclusion

Within this paper, we have defined the mechanisms needed to secure voice traffic over wireless and IP network. Different architectures available related to ciphering, authentication and integrity mechanisms were presented. Some security mechanisms are applied at the access level leaving the end-to-end communication unsecured (GSM network). Bluetooth authenticates the device only while encrypting the transferred data. IPsec provides different security protocols introducing more complexity and resource usage. A new profile of the Real-time Transport Protocol which provides security for the voice packets was explored. The ITU-T organization defined within the H.235 recommendation the security needs to be incorporated within its voice over IP standards. SIP uses different security schemes based on TLS/SSL or HTTP-S. The intent of this paper is to outline the different security techniques implemented in wireless and IP network. User authentication was proposed based on smart cards providing also non-repudiation functionality, while integrity, confidentiality and replay protection were provided by SRTP.
We are implementing and developing a prototype that integrates the proposed security solution with the current VoIP standards which will help us testing and evaluating performance criteria.

9. References

[1] S. Kent, R. Atkinson, *"Security Architecture for the Internet Protocol", IETF RFC 2401* - November 1998
[2] C. Brookson, *"GSM (and PCN) Security and Encryption"*, http://gsmsecurity.com/papers.shtml
[3] X. Lagrange, P. Godlewski, S. Tabbane, *"Réseaux*

GSM", Hermes Sciences, 5th edition
[4] R. Nichols, P. Lekkas, "*Wireless Security, Models, Threats and Solutions"*, McGraw-Hill Telecom Professional edition – 2002
[5] T. Dierks, C. Allen, "T*he TLS Protocol Version 1.0", IETF RFC 2246*, January 1999
[6] ITU-T, "*Packet-based multimedia communications systems"* Recommendation H.323 version 4
[7] ITU-T, "*Security and encryption for H-Series multimedia terminals"*, Draft Recommendation H.235 - version 3
[8] ITU-T Recommendation H.235, "*Annex D - Baseline Security Profile"*
[9] ITU-T Recommendation H.235, "*Annex E – Signature Profile"*
[10] ITU-T Recommendation H.235, "*Annex F - Hybrid Security Profile"*
[11] ITU Recommendation H.235, "*Annex I – H.323 Implementation Details"*
[12] R. Morrow, "*Bluetooth: Operation and Use"*, McGraw-Hill Professional, 1st edition 2002
[13] S. Kent, R. Atkinson, "*IP Authentication Header", IETF RFC 2402*, November 1998
[14] S. Kent, R. Atkinson, "*IP Encapsulating Security Payload (ESP)", IETF RFC 2406*, November 1998
[15] R. Barbieri, D. Brushi, E. Rosti, "*Voice over IPsec: Analysis and Solutions"*, 18th Annual Computer Security Applications Conference December 2002, San Diego California
[16] M. Baugher, D. McGrew, E. Carrara, M. Naslund, K. Norrman, "*The Secure Real-time Transport Protocol SRTP", http://www.ietf.org/internet-drafts/draft-ietf-avt-srtp-09.txt*, July 2003
[17] **"***IEEE 802.11"*. Available from http://grouper.ieee.org/groups/802/11/main.html
[18] M. Michels, "*Designing VoIP Networks: Lessons from the Edge"*, Business Communications review/ 2003
[19] A. Freier, P. Karlton, P. Kocher, "*The SSL Protocol Version 3.0", ,draft-freier-ssl-version3-02.txt*, November 18, 1996
[20] M. Handley, H. Schulzrinne, E. Schooler, J. Rosenberg, "*SIP: Session Initiation Protocol", IETF RFC 2543* - March 1999
[21] S. Frankel, R. Glenn, S. Kelly, "*The AES-CBC Cipher Algorithm and Its Use with IPsec", IETF RFC 3602* – September 2003
[22] L. Pesonen, "*GSM Interception*", Helsinki University of Technology, November 1999
[23] NIST, FIPS Publica tion 46-3, "*Data Encryption Standard (DES)"*, October 1999.
[24] Greg Rose, "*A precis of the new attacks on GSM encry*ption", QUALCOMM Australia, September 2003.
[25] M. Handley, V. Jacobson, "*SDP: Session Description Protocol", IETF RFC 2327*, April 1998
[26] D. Koch, "*SIP Security CS 756M Project"*, University of Waterloo, August 2001
[27] D. Atkins, W. Stallings, P. Zimmermann, "*PGP Message Exchange Formats"*, IETF RFC 1991, August 1996
[28] B. Ramsdell, "*S/MIME Version 3 Message Specification*", IETF RFC 2633 June 1999
[29] **"***Security in ISDN – Chapter 4"*, Available at http://csrc.nist.goc/publications/nistpubs/500-189/isdn4.ps
[30] H. Schulzrinne, S. Casner, R. Frederick, V. Jacobson, "*RTP: A transport Protocol for Real-Time Applications", draft-ietf-avt-rtp-news-12.ps*, March 2003
[31] NIST, "*Advanced Encryption Standard (AES)"*, FIPS PUB 1997, http://www.nist.gov/aes
[32] H. Lipmaa, P. Rogaway, D. Wagner, "*CTR-Mode Encryption"*, NIST, http://csrc.nist.gov/encryption/modes/workshop1/papers/lipmaa-ctr.pdf
[33] 3GPP TS 35.201 V4.1.0 (2001-12) Technical Specification 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; 3G Security; Specification of the 3GPP Confidentiality and Integrity Algorithms; Document 1: f8 and f9 Specification(Release 4).
[34] 3GPP TR 33.908 V4.0.0 (2001-09) Technical Report 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; 3G Security; General Report on the Design, Specification and Evaluation of 3GPP Standard Confidentiality and Integrity Algorithms (Release 4).
[35] K. Krawczyk, M. Bellare, R. Canetti, "*HMAC: Keyed-Hashing for Message Authentication"*, IETF RFC 2104, February 1997
[36] R. Rivest, *"The MD5 Message-Digest Algorithm"*, IETF RFC 1321, April 1992
[37] C. Madson, N. Doraswamy, *"The ESP DES-CBC Cipher Algorithm with Explicit IV"*, IETF RFC 2405, November 1998.
[38] NIST, FIPS PUB 180-1, "*Secure Hash Standard"*, April 1995