A 3G IMS-BASED TESTBED FOR SECURE REAL-TIME AUDIO SESSIONS

Paolo Cennamo, Antonio Fresa, Anton Luca Robustelli, Francesco Toro
Co.Ri.TeL, Via Ponte Don Melillo, I-84084 Fisciano (SA), Italy

Maurizio Longo, Fabio Postiglione
Dipartimento di Ingegneria dell'Informazione ed Ingegneria Elettrica
Università degli Studi di Salerno, Via Ponte Don Melillo, I-84084 Fisciano (SA), Italy

Keywords: Beyond-3G networks, IP Multimedia Subsystem, Secure Real-Time Protocol, multimedia communications, voice quality.

Abstract: The emerging all-IP mobile network infrastructures based on 3rd Generation IP Multimedia Subsystem philosophy are characterised by radio access technology independence and ubiquitous connectivity for mobile users. Currently, great focus is being devoted to security issues since most of the security threats presently affecting the public Internet domain, and the upcoming ones as well, are going to be suffered by mobile users in the years to come. While a great deal of research activity, together with standardisation efforts and experimentations, is carried out on mechanisms for signalling protection, very few integrated frameworks for real-time multimedia data protection have been proposed in a context of IP Multimedia Subsystem, and even fewer experimental results based on testbeds are available. In this paper, after a general overview of the security issues arising in an advanced IP Multimedia Subsystem scenario, a comprehensive infrastructure for real-time multimedia data protection, based on the adoption of the Secure Real-Time Protocol, is proposed; then, the development of a testbed incorporating such functionalities, including mechanisms for key management and cryptographic context transfer, and allowing the setup of Secure Real-Time Protocol sessions is presented; finally, experimental results are provided together with quantitative assessments and comparisons of system performances for audio sessions with and without the adoption of the Secure Real-Time Protocol framework.

1 INTRODUCTION

The very rapid evolution of the communication infrastructures has progressively rendered access to communication facilities ubiquitous.

These fast-evolving communication technologies have greatly stimulated research activity on security issues, encompassing both data confidentiality and data protection, both in corporate and residential environments.

While, on the one hand, the legacy mobile digital networks (GSM, GPRS, UMTS) provide strong security and confidentiality guarantees, on the other, the emergence of the 3rd Generation IP Multimedia Subsystem (IMS) as the unified and standard platform, based on the *all-IP* paradigm for the provision of real-time multimedia services both to mobile and fixed users, is bringing the security issues to the forefront once again. Indeed, the adoption for next-generation mobile networks of an IP-based transport infrastructure, based on Internet Engineering Task Force (IETF) protocols, both for signalling, based on Session Initiation Protocol (SIP) (Rosenberg et al., 2002), and multimedia real-time data transport, by using Real-time Transport Protocol (RTP) (Schulzrinne et al., 2003), will expose future mobile telecommunication infrastructures to all the security threats (and maybe new ones) of the public Internet. This emerging scenario requires specific research and definitions of solutions aiming to guarantee acceptable levels of user data confidentiality and protection.

Another key feature of future IMS-based networks will be the access-domain independence, i.e. the IMS service provision infrastructure will be totally independent of the particular radio technologies deployed in the access network. In other words, future IMS will be *access-agnostic* and it will work on the top of any kind of wired or wireless access technology;

Cennamo P., Fresa A., Luca Robustelli A., Toro F., Longo M. and Postiglione F. (2007).
A 3G IMS-BASED TESTBED FOR SECURE REAL-TIME AUDIO SESSIONS.
In *Proceedings of the Second International Conference on Security and Cryptography*, pages 125-132
DOI: 10.5220/0002118601250132

on the other hand, given that each access technology provides different security guarantees (varying from very strong to none at all), the IMS cannot in general rely on such capabilities. Then, IMS-specific security mechanisms must be provided, and such mechanisms can only operate from the IP Layer upwards (network, transport or application), being IP the first common technological layer envisaged by the IMS philosophy.

The paper is organized as follows. First of all, we point out the advantages of adopting appropriate security protocols in order to protect both the signalling and the real-time multimedia flows, for which we focus on the *Secure RTP protocol* (SRTP) (Baugher et al., 2004); then, we propose a architectural solution to involve security mechanisms during sessions establishment and control and we describe the developed testbed which implements it within an IMS-like prototype. Finally, we provide some quantitative evaluations and comparative assessments related to voice quality parameters measured in end-to-end audio sessions, pointing out the the influence of the SRTP framework deployment on voice communication quality.

2 IMS: ARCHITECTURE AND SECURITY ISSUES

Most researchers consider IMS as the key element in the next generation network architectures since it enables the convergence of data, speech, and mobile network technologies over a unified IP-based infrastructure. The organization responsible for the definition of Beyond-3G (also known as B3G) mobile communication systems, including IMS, is the Third Generation Partnership Project (3GPP) (3GPP, The 3rd Generation Partnership Project, 1998). The 3GPP has chosen SIP as the signalling protocol for the setup, modification and tear-down of multimedia sessions. The Call Session Control Function (CSCF) servers represent the core elements, within the IMS, for the management of the SIP signalling. The Proxy CSCF (P-CSCF), usually located in the Visited Network, represents the first contact point for the user terminal towards the IMS network and takes care of forwarding the SIP signalling towards the subscriber's Home Network; the Serving CSCF (S-CSCF) is probably the main CSCF server and is located in the subscriber's Home Network (typically the operator to which the user is subscribed): its task is to process the SIP signalling, take decisions on managing the multimedia sessions. Another important function of the IMS architecture is the Home Subscriber Server (HSS) database that contains all the user-related subscription data required to handle a multimedia session, such as information on user location, security data and user profiles. The interaction among the three CSCF nodes and the HSS allows the complete management of the SIP signalling necessary for the establishment and support of the multimedia sessions.

Nowadays, millions of customers are using computer networks for e-banking, e-commerce and submitting their tax returns and since 3G architecture aims to enable such secure transactions together with real-time services in its IP-based infrastructure, the security issues have acquired a primary importance. Network security problems can be roughly divided into six closely related areas, each of them with its peculiar goals:

- *Authentication*: to guarantee user identity;
- *Confidentiality*: to keep information out of the hands of unauthorised users;
- *Integrity*: to avoid information alteration or the whole substitution of messages by malicious users;
- *Non-repudiation*: to avoid that users deny having sent or received information actually sent or received by them;
- *Authorization*: to allow only authorised users to access particular resources and services;
- *Availability*: to guarantee the effectiveness of a service avoiding actions of disturbance by malicious users.

There are many possible approaches in order to provide security services; indeed, security features can be implemented in different layers of the TCP/IP reference stack: at the Network Layer by adopting IPsec (Thayer et al., 1998), at the Transport Layer by TLS (Dierks and Allen, 1999) and at the Application Layer using HTTP Digest (Franks et al., 1999) or other.

Security in an IMS scenario can be categorized as follows (Koien, 2002):

- *Access security*: it includes mutual authentication, encryption and integrity of both signalling and multimedia data which are exchanged between the B3G terminal and the network;
- *Network security*: it deals with traffic protection between network nodes, which can belong to the same operator or different ones.

The IMS adopts IPsec for signalling protection both in the access and network domains but nothing is specified for data or multimedia traffic. In order to accommodate all requirements (very different and

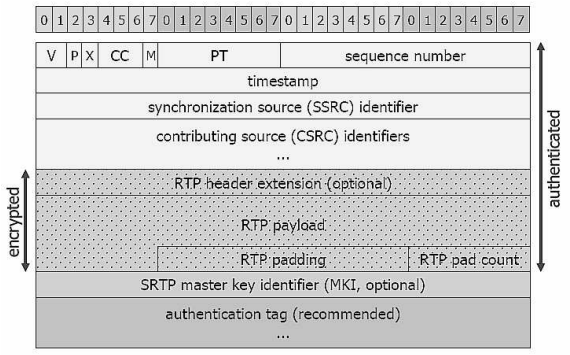

Figure 1: A SRTP message.

sometimes very stringent) of data and real-time applications, security protocols adapted to the single application (and thus working at the Application Layer) appear more appealing.

In particular, four factors must be taken into account in order to protect multimedia real-time communications: bandwidth availability, delay, computational power of the mobile terminals and transmission-error sensitivity. To address these issues a very promising choice is the adoption of the Secure Real-time Transport Protocol (SRTP) (Baugher et al., 2004). This protocol employs particular transforms such as the Advanced Encryption Standard (AES) symmetric cipher (Schaad and Housley, 2002), since symmetric cryptography is characterised by lower delays and computational burden with respect to asymmetric cryptography (Stallings, 2004); furthermore, such a ciphering system can avoid the error propagation drawback if used in a *stream* modality. Another advantage of SRTP is related to bandwidth consumption due to the frequent re-keying procedures occurring during long real-time sessions. In fact, it introduces a *32-bit RollOver Counter* (ROC) in order to expand the space of the RTP sequence numbers and eliminate the need of re-keying (Blom et al., 2002).

Alternative proposals are available for multimedia real-time protection based on IP tunneling protocols, such as IPsec, which seem to suffer some performance limitations (Ranganathan and Kilmartin, 2001; Vaidya et al., 2005).

3 THE SECURE REAL-TIME TRANSPORT PROTOCOL FRAMEWORK

The SRTP protocol was designed to be deployed in heterogeneous network architectures; the critical factors which were taken into consideration were bandwidth, delay, the need of computational resources and transmission errors. SRTP is an RTP profile and it can be considered as a sub-layer implementation located between the RTP application protocol and the transport protocol: on the sending side, SRTP first intercepts RTP packets and then forwards equivalent SRTP packets; on the receiving side, it first intercepts SRTP packets and then relays equivalent RTP packets upwards. On the other hand, the Real-time Transport Control Protocol (RTCP) is secured by Secure RTCP (SRTCP) so as SRTP does to RTP. Message authentication based on SRTCP is mandatory when SRTP is used; moreover, it can protect the RTCP fields to keep track of session members, it can provide feedbacks to RTP senders and securely manage counters of packet sequence.

Then, SRTP provides a framework for authentication and encryption of RTP and RTCP data streams. Specifically, SRTP proposes a set of default cryptographic algorithms and it also allows for the introduction of new ones in the future. Together with appropriate mechanisms for key management, SRTP can effectively provide security services to RTP applications both of unicast and multicast transmissions.

Fig. 1 depicts the format of a SRTP message. The specific additional fields introduced by SRTP are:

- *Master Key Identifier (MKI)*: the key management mechanism defines and uses this field. MKI identifies the master key from which the session keys can then be derived. Authentication and/or encryption of the RTP packets is then performed by using such session keys.
- *Authentication Tag*: this field is employed in order to carry message authentication data. The *Authenticated Portion* of an SRTP packet is made up of the RTP header followed by the *Encrypted Portion* of the SRTP packet. If both encryption and authentication are applied, encryption must be applied before authentication on the sending side and vice-versa on the receiving side. The Authentication Tag provides authentication of the RTP header and payload, and it also provides, even if indirectly, replay protection by authenticating the sequence number. It is worth noting that the MKI is not integrity-protected since this would provide no additional protection.

3.1 The Cryptographic Context

Each SRTP stream requires the sender and the receiver to maintain cryptographic state information; moreover, in order to establish an SRTP session two

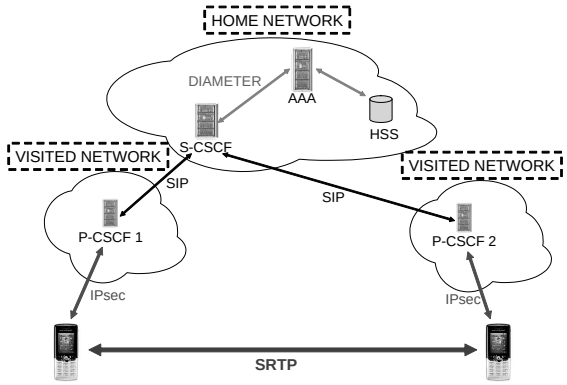

Figure 2: The proposed IMS-based architecture for secure multimedia communications.

users have to come to an agreement on specific parameters, such as the cryptographic and integrity techniques to use, the master key from which the session keys will be derived and so on. All this information is called *Cryptographic Context* and is handled by a key management mechanism external to SRTP. Several key management standards have been proposed for SRTP cryptographic contexts, such as MIKEY (Arkko et al., 2004) and KEYMGT (Arkko et al., 2006).

4 THE PROPOSED SRTP-BASED COMMUNICATION SYSTEM

In this Section we present the architecture of a Beyond-3G network built by integrating the SRTP framework into the IMS infrastructure. The adoption of SRTP to protect real-time communications is motivated by performance limitations in IPsec-based solutions, as already stated in Sect. 2. Furthermore, SRTP allows to propose a novel technique for the key exchange mechanism, based on IMS SIP signalling, which does not require any additional message thus increasing the overall performance (see Sect. 4.2).

4.1 The IMS-based Testbed Architecture

As previously sketched, with the present work we aim to show how it is possible to integrate a secure architecture for multimedia communications into a B3G network scenario. In order to introduce the several actors involved in such a scenario, it might prove useful to briefly illustrate the environment we adopted for the developed testbed.

Our IMS-like prototype is represented in Fig. 2, where two mobile phones are included which are connected to the IMS Home Network (i.e. with the S-CSCF server) through the P-CSCF nodes of each mobile user's Visited Network. The Authentication, Authorisation, Accounting (AAA) server (Senatore et al., 2004) is introduced according to the IMS architecture defined by 3GPP. In the depicted network scenario, when a user switches his mobile phone on, a registration phase takes place by means of the AKA protocol (AKA, 2003) encapsulated within SIP REGISTER messages. This protocol allows a mutual authentication (that is the user and the network authenticate each other): through this procedure each user independently computes the cryptographic and integrity keys which will be used in subsequent secure communications. Furthermore, during the registration phase an IPsec security association is established between each user and its reference P-CSCF in order to guarantee a strong protection for the subsequent SIP signalling messages.

4.2 A IMS-based Master Key Exchange Mechanism

RFC 3711 provides two different methods for selecting the Master Key that can be used during an SRTP session: the first mechanism proposes the use of the MKI of the SRTP packet header, while the second provides the definition of a *(From, To)* mechanism, as already explained in Subsect. 3.

In our IMS testbed, we adopt a novel mechanism for cryptographic context transfer which does not introduce either a new protocol or a new messages exchange, which would be both a burden to the signalling system and a cause of delay; in fact, all of the information necessary to establish the SRTP session can be encapsulated within the SIP signalling messages, in appropriate fields of the Session description Protocol (SDP) (Handley and Jacobson, 1998) body, already conveyed by SIP messages for multimedia sessions.

In particular, we integrated into our B3G network prototype a mechanism to transfer the cryptographic context based on the encapsulation of context information into the SIP INVITE transaction. Fig. 3 shows the whole signalling process taking place between two end users for a multimedia session set-up.

An important role is played by the S-CSCF and the AAA server; indeed, by parsing the SIP INVITE body the S-CSCF detects the presence of the cryptographic context attribute in the SIP message and consequently sends a specific request to the AAA server,

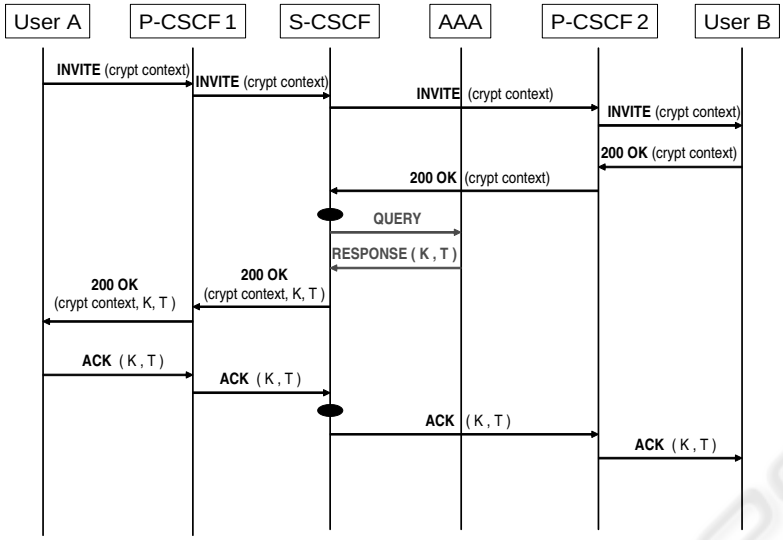

Figure 3: A SIP-based Master Key exchange mechanism.

which answers back by a message with the session key and its lifetime for the under way SRTP session. Then, the S-CSCF includes these parameters into the 200 OK message and subsequent ACK message concluding the INVITE transaction. At this point, the two users are able to establish an SRTP session. When the timer expires, a *re-INVITE* message is sent between the users in order to re-negotiate the cryptographic context and the previously described procedure takes place once again.

4.3 A SRTP-enabled Voice Application

Besides the IMS-based solution, in this paper we propose and develop a Voice over IP (VoIP) application that implements the SRTP framework by modifying the open-source Robust Audio Tool (RAT) code, version 4.2.25 (Robust Audio Tool (RAT), 2004). In the application we propose for secure real-time communications, it is possible to distinguish two different modules:

- the SIP module for the signalling and session control;
- the RAT module for the audio communication management.

These two modules need to exchange information using a particular communication channel: in our solution, such modules are organized according to the schema reported in Fig. 4.

The SIP module receives the Master Key within the SIP signalling flow, as described in Sect. 4.2. By means of a local *Message Bus* (Mbus), such a key is then transferred to the *RAT Controller* which forwards it to the *Media Engine*: thus, this Master Key becomes the Active Key for the actual SRTP session.

Let us recall that the SRTP layer is located, within the protocol stack, between the Transport Layer (in this case, UDP) and the Application Layer (RTP).

It is worth pointing out that, during the definition phase, we evaluated two different approaches to the integration of SRTP in the VoIP application: the former aimed to maintain RTP and SRTP as separate modules, the latter to create a hybrid RTP/SRTP structure. At the end, we decided upon the latter solution since it requires less computational burden and achieves better performances, which are crucial constraints for an application that has to process real-time media.

4.3.1 The SRTP Transmission Phase

Our implementation does not modify the RTP packing phase and operates on the packet which is ready to be sent. The packet is "intercepted" inside the `rtp_send_data()` function and, if the session encryption is enabled, the *SRTP packet setting phase* starts.

The SRTP packet setting phases implemented in our prototype are those provided by RFC 3711. The first phase consists of the search of the active key by using the `find_key()` method: this phase is strictly related to the exchange key mechanism described in Sect. 4.2 and uses the key exchanged during the SIP session setup. The second phase concerns the generation of the session keys by means of the `key_derivation()` method through the ac-

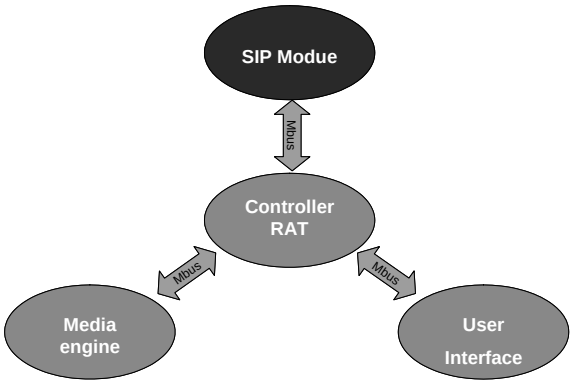

Figure 4: A schematic representation of the developed Voice over IP application.

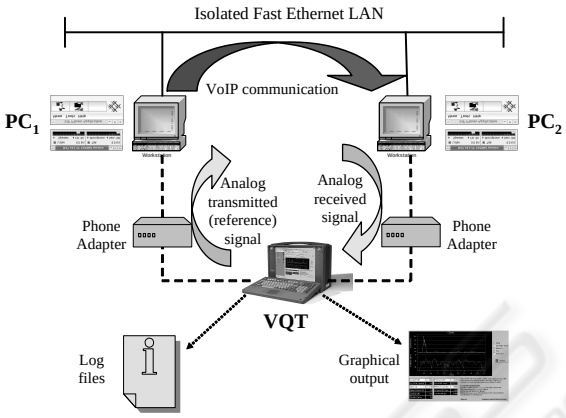

Figure 5: The voice quality measurement testbed.

tive key obtained in the first stage. The third phase schedules the keystream generation by invoking a `keystream_generator()` function. In the fourth phase, the payload is encoded by performing a simple XOR operation between the generated keystream and the payload itself. The resulting encrypted payload replaces the non-encrypted RTP packet payload, since their length are the same. The last phase schedules the *Authentication Tag* generation by means of the `hmac_SHA1()` function. This tag is appended to the RTP packet so that the receiver can authenticate the packet itself by it.

When those phases are accomplished, the Controller switches back to the `rtp_send_data()` function, that delivers the packet thus prepared to the `udp_send()` function, which finally sends the UDP segment toward the receiver.

4.3.2 The SRTP Reception Phase

Similarly to what happens at the beginning of the transmission phase, the received packet is intercepted within `rtp_receive_data()` function and the SRTP session management starts.

The first action to be performed is finding the active key by using the `find_key()` method, as described for the transmission phase. This key is then passed to the `key_derivation()` method in order to generate the session key. At this stage the `hmac_SHA1()` function locally generates the *Authentication Tag* to be compared with the one received within the SRTP packet. If the two tags match, the received packet is authentic and it is thus possible to go on with the decryption process. The next step is devoted to the estimation of the ROC related to the *Sequence Number* of the RTP packet. This process takes place after the packet authentication procedure and it is needed in order to estimate the correct counter value of the SRTP packet.

The `keystream_generator()` method generates the keystream that is used in the `dencript_payload()` method to execute decryption. Subsequently, the control switches back to the `rtp_receive_data()` method and the audio content can be reproduced by the application as usual.

In line of principle, the introduction of the SRTP framework introduces, both at the transmission and reception side, an increase of complexity and computational load that might potentially badly influence the system performances. That is why it appears particularly worthwhile to analyse the impacts of our SRTP implementation on a real end-to-end audio communication, as reported in the following section.

5 VOICE QUALITY EVALUATION

In order to assess the influence of our SRTP implementation on real VoIP communications, we compare quantitatively system performances in terms of both the mean audio delay and the mean perceived voice quality at the receiver.

The measurement testbed is shown in Fig. 5, where the VoIP applications involved in the audio communications are connected to the same (isolated) Fast Ethernet LAN segment (no signalling server is involved). In such an environment, the delay introduced by the network can be considered negligible. Measurements are collected by an Agilent Voice Quality Tester (VQT) connected, by means of proper Phone Adapters, to the audio card line-in of PC_1 and to the audio card line-out of PC_2 in the same system

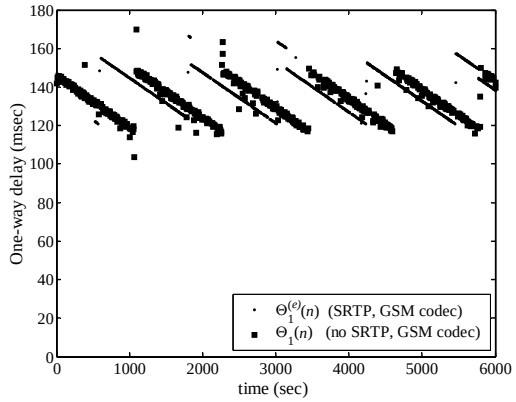

Figure 6: Typical one-way delay time series provided by VQT for application running on PCs not specialized for real-time applications. The adopted codec is GSM.

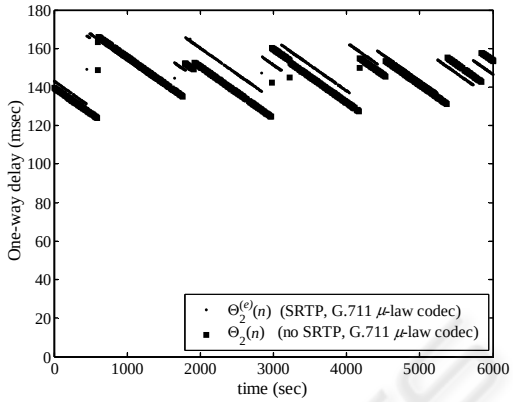

Figure 7: Typical one-way delay time series provided by VQT for application running on PCs not specialized for real-time applications. The adopted codec is G.711 μ-law.

configuration during audio sessions both in case VoIP applications implement the SRTP framework, as described in Sect. 4, and in case of no SRTP implementation. The VQT can represent measurements results both in a textual (log files) and a graphical way.

PC_1 and PC_2, where VoIP applications are running, are both Linux-based machine with a *Pentium IV* 2.8 GHz processor and audio cards, whose quality can heavily influence quality measurements, are both *Creative SoundBlaster Live!*.

We select two audio codec available in the VoIP applications under test: (full-rate) GSM and G.711 implementing the companding μ-law algorithm (Bellamy, 2000). Every transmitted packet contains 20 msec of speech, i.e. the RTP payload is 33 byte or 160 byte long for GSM and G.711, respectively. The encryption algorithm used by SRTP in our measurements is *AES* in Counter Mode (AES-CTR).

A key parameter that influences the quality of real-time communication is the one-way delay experienced by two speakers during an audio session, also known as *mouth-to-ear* delay (Jiang et al., 2003). In order to assess the impact of SRTP on the total delay, we collect $N = 1000$ measures of audio delay for each session (one using SRTP and one without SRTP for each audio codec), where every delay sample $\Theta(n)$ for $n = 1, ..., N$ is computed every $\Delta = 6$ sec by the VQT by estimating the position of the maximum of the cross-correlation between the transmitted signal and the received one. The measured mouth-to-ear delay time series for GSM and G.711 μ-law are indicated as $\Theta_1(n)$ and $\Theta_2(n)$, respectively, while the presence of SRTP is pointed out by the superscript (e). Typical delay time series (with and without SRTP sublayer) are shown in Figs. 6 and 7 for VoIP applications running on operating systems not specialized for real-time applications, such as Linux or Windows, where general purpose schedulers can cause some artifacts, such as a piecewise linear decrease (or increase) of the one-way delay.

The computed average delays, reported in Table 1, seem to indicate that AES-CTR encryption has no significant influence (just few milliseconds).

Another key parameter to assess audio communication quality is the perceived quality of the speech at destination. A widely adopted tool for objective measurements of it is the *Perceptual Evaluation of Speech Quality* (PESQ), described in ITU-T Rec. P.862 (Beerends et al., 2002), which uses a sensory model to compare the transmitted signal with the receiving one. In order to relate its results to the traditional subjective quality score *Mean Opinion Score* (MOS), based on time-consuming human listeners interviews, the *PESQ Listening Quality* (PESQ-LQ) is often used, providing values ranging from 1 (bad) to 4.5 (very good).

The average PESQ-LQ values were computed on 30 voice clarity measurements, provided again by the VQT using English speech samples, and are reported in Table 2, where it is possible to notice that the SRTP does not introduce any appreciable variation on the quality of the speech. This agrees with the general conception that encryption in itself should not cause

Table 1: Average mouth-to-ear delays (in msec).

	GSM	**G.711 μ-law**
no SRTP	132.71	144.14
SRTP	138.86	149.79

Table 2: Average PESQ-LQ.

	GSM	**G.711 μ-law**
no SRTP	3.693	4.028
SRTP	3.690	4.013

an information loss.

Summing up, the introduction of the SRTP framework does not seem to influence speech quality from a practical prospective in our prototype testbed. Then, SRTP seems suitable for a deployment in real-world network scenario.

6 CONCLUSIONS

The introduction of real-time cryptography technique for the multimedia flows with the adoption of the SRTP protocol is aimed to guarantee a good security level to multimedia communications.

One of the mechanism that offer a good level of robustness against the *two time pad* typologies of attacks is the introduction of a periodic key update mechanism. In our proposal the update mechanism towards IMS SIP signalling does not introduce any increase of the number of exchanged messages, as it may happen adopting a Master Key Identifier or a *(From,To)* mechanism.

The quality of the communication does not turn out to be degraded even though a real-time cryptography and de-cryptography is performed. In particular, the SRTP framework does not influence the quality of the speech during VoIP communications, both in terms of delay and PESQ-LQ index.

Future developments will concern, first of all, the practical establishment of an SRTP session also for the video content between two users within the IMS-like prototype. Another development will be related to the adaptation of our architectural solution to a multi-conferencing scenario.

REFERENCES

3GPP, The 3rd Generation Partnership Project (1998). http://www.3gpp.org/.

AKA (2003). Authentication and key agreement. 3GPP TS 33.102 version 6.0.0.

Arkko, J. et al. (2004). MIKEY: Multimedia internet keying. IETF RFC 3830, http://www.ietf.org/rfc/rfc3830.txt.

Arkko, J. et al. (2006). Key management extension for session description protocol (SDP) and real time streaming protocol (RTSP). IETF RFC 4567, http://www.ietf.org/rfc/rfc4567.txt.

Baugher, M. et al. (2004). The secure real-time transport protocol (SRTP). IETF RFC 3711, http://www.ietf.org/rfc/rfc3711.txt.

Beerends, J., Hekstra, A. P., Rix, A. W., and Hollier, M. P. (2002). Perceptual evaluation of speech quality (PESQ), the new ITU standard for end-to-end speech quality assesment, part i & ii. 50(10):755–778.

Bellamy, J. (2000). *Digital Telephony*. Wiley-Interscience, 3rd edition.

Blom, R., Carrara, E., Lindholm, F., Norman, K., and Naslund, M. (2002). Conversational IP multimedia security. In *Proc. 4th IEEE MWCN 2002*, pages 147–151.

Dierks, T. and Allen, C. (1999). The TLS protocol. IETF RFC 2246, http://www.ietf.org/rfc/rfc2246.txt.

Franks, J. et al. (1999). HTTP authentication: Basic and digest access authentication. IETF RFC 2617, http://www.ietf.org/rfc/rfc2617.txt.

Handley, M. and Jacobson, V. (1998). SDP: Session description protocol. IETF RFC 2327, http://www.ietf.org/rfc/rfc2327.txt.

Jiang, W., Koguchi, K., and Schulzrinne, H. (2003). QoS evaluation of VoIP end-points. In *Proc. IEEE ICC 2003*, volume 3, pages 1917–1921.

Koien, G. M. (2002). An evolved UMTS network domain security architecture. Technical report, R&D Telenor.

Ranganathan, M. K. and Kilmartin, L. (2001). Investigations into the impact of key exchange mechanisms for security protocols in VoIP networks. In *Proc. First Joint IEI/IEE Symposium on Telecommunications Systems Research*. http://telecoms.eeng.dcu.ie/symposium/papers/D2.pdf.

Robust Audio Tool (RAT) (2004). http://www-mice.cs.ucl.ac.uk/multimedia/software/rat/.

Rosenberg, J. D. et al. (2002). Session Initiation Protocol (SIP). IETF RFC 3261, http://www.ietf.org/rfc/rfc3261.txt.

Schaad, J. and Housley, R. (2002). Advanced encryption standard (AES) key wrap algorithm. IETF RFC 3394, http://www.ietf.org/rfc/rfc3394.txt.

Schulzrinne, H. et al. (2003). RTP: A transport protocol for real-time applications. IETF RFC 3550, http://www.ietf.org/rfc/rfc3550.txt.

Senatore, A., Fresa, A., Robustelli, A. L., and Longo, M. (2004). A security architecture for access to the IP multimedia subsystem in B3G networks. In *Proc. 7th WPMC 2004*.

Stallings, W. (2004). *Data and Computer Communications*. Prentice Hall, 7th edition.

Thayer, M. et al. (1998). IP security document roadmap. IETF RFC 2411, http://www.ietf.org/rfc/rfc2411.txt.

Vaidya, B., Kim, J., Pyun, J., Park, J., and Han, S. (2005). Performance analysis of audio streaming in secure wireless access network. In *Proc. 4th IEEE ACIS 2005*, pages 556–561.