Paper: ASAT-13-CE-13

13th International Conference on
AEROSPACE SCIENCES & AVIATION TECHNOLOGY,
ASAT- 13, May 26 – 28, 2009, E-Mail: asat@mtc.edu.eg
Military Technical College, Kobry Elkobbah, Cairo, Egypt
Tel : +(202) 24025292 – 24036138, Fax: +(202) 22621908

Performance Analysis of Transmitting Voice over Communication Links Implementing IPsec

Gouda I.Salama*, M. Elemam Shehab*, A. A. Hafez*, M. Zaki**

Abstract: The prevalence and ease of packet sniffing and other techniques for capturing packets on an IP based network makes encryption a necessity for VOIP(Voice Over Internet Protocol). Security in VOIP is concerned both with protecting what a person says(by Encryption) as well as to whom the person is speaking(by Authentication). IPsec can be used to achieve both of these goals as long as it is applied with ESP(Encapsulating Security Payload) using the tunnel method. This secures the identities of both the endpoints and protects the voice data from prohibited users once packets leave the corporate intranet. The incorporation of IPsec with IPv4 increase the availability of encryption, VOIPsec (VOIP using IPsec) helps reduce the threat of man in the middle attacks, packet sniffers, and many types of voice traffic analysis. Combined with the firewall implementations, IPsec makes VOIP more secure than a standard phone line, In this paper the negative effect of adding security to VoIP networks has been measured for different simulation times using OPNET(Network Simulation Tool). the results show that transmitting voice over IPsec increase the end to end delay,delay variation(jitter),packet loss and call setup time .

Keywords: VOIP, SIP, IPsec

1. Introduction

In recent years, we have witnessed a growing interest in the transmission of voice using the packet based protocols. Voice over Internet protocol (VoIP) is a rapidly growing technology that enables the transport of voice over data networks such as the public Internet the following steps are performed to carry the voice signal over IP based network: [1]

- Digitization of the analog signal;
- Packet generation of the digital signal according to the TCP-UDP/IP protocols;
- Transmission of the packets on the network;
- Packet reception and analog signal reconstruction at the destination.

When sending voice traffic over IP networks, a number of factors contribute to overall voice quality as perceived by an end user. The factors determine voice quality include the choice of codec, echo control, packet loss, delay, delay variation (jitter), and the design of the network.

* Egyptian Armed Forces
** Professor,. Azhar university, Cairo, Egypt, melemam @hotmail.com

If the End to End delay becomes too long, the conversation begins to sound like two parties talking on a Citizens Band radio.A buffer in the receiving device tries to compensate for jitter (delay variation). If the delay variation exceeds the size of the jitter buffer, there will be buffer overruns at the receiving end, with the same effect as packet loss anywhere else in the transmission path. The incorporation of IPsec with IPv4 increase the availability of encryption, VOIPsec (VOIP using IPsec) helps reduce the threat of man in the middle attacks, packet sniffers, and many types of voice traffic analysis. Combined with the firewall implementations, IPsec makes VOIP more secure than a standard phone line, In this paper the negative effect of adding security to VoIP networks has been measured for different simulation times using OPNET simulator. our results show that transmitting voice over IPsec increase end to end delay, delay variation(jitter),packet loss and call setup time.

There are many methods for header compression that are defined [2,3,4], but the general principles of operation are very similar and essentially comprise the following elements:

- The full header is sent with the first datagram of the communication and stored by the receiver;
- Each field can be classified as UNCHANGING, RANDOM changes, DELTA changes or inferred as DEFAULT;
- The only segments of header information that need to be sent in every header are fields that change often and randomly, such as checksums or authentication data;
- For fields that are incremented from the previous value (DELTA), only the delta increment is sent.

An approach to the QOS issues associated with VOIPsec is proposed by Barbieri et al. [2] at the conclusion of their study of VOIPsec traffic. Their solution targets the increase of packet size stemming from the use of IPsec .they implemented cIPsec: a version of IPsec that compresses the internal header of a packet down to approximately four bytes .this is possible because much of the data in the internal headers of a packet remains constant or is duplicated in the outer header. one thing they didn't consider is the actual time required to perform the compression may take much longer than the time saved in crypto-engine it is also important to note that the compression scheme used in cIPsec only compresses the packet header. the compression QOS issues associated with codec are not applicable in this scenario because no actual media is being considered, only the IP headers. however QOS is changing according to the change in codec and for more compression.

The factors determine voice quality include the choice of codec, echo control, packet loss, delay, delay variation (jitter), and the design of the network, Delay and jitter are two of the most critical factors that affect the quality of audio transmission, This paper is focused on showing the negative effect of adding internet protocol security to VOIP applications and we will specially focused on the delay , delay variation , packet loss and call setup time.
This remained paper will organize as follows; section 2 presents a quick overview for used protocols, section 3 presents VoIP security requirements and quality metrics, section4 presents performance analysis and section 5 presents conclusion.

2. Protocols

Currently, there is no single VoIP signaling protocol, which has been exclusively adopted by the networking community. However, it is widely accepted that the Session Initiation Protocol (SIP) has a number of distinct advantages over H.323 and MEGACO, most notably its simplicity[7]. Additionally, the IPSec framework of security protocols and architectures offer a myriad of flexible options when it comes to secure a VoIP network. For this reason, SIP and IPSec were chosen as the signaling and security protocols/architecture of choice on which the simulation model for the secure VoIP network was based. The simulation model was developed using the OPNET Modeler (Network Simulation Tool).

2.1 Session Initiation Protocol

Session Initiation Protocol (SIP)[5] is the Internet Engineering Task Force's (IETF's) standard for multimedia conferencing over IP. SIP is an ASCII-based application-layer control protocol that can be used to establish, maintain, and terminate calls between two or more end points. Like other VoIP protocols, SIP is designed to address the functions of signaling and session management within a packet telephony network. Signaling allows call information to be carried across network boundaries. Session management provides the ability to control the attributes of an end-to-end call. A motivating goal for SIP was to provide a signaling and call setup protocol for IP-based communications that can support a superset of the call processing functions and features present in the public switched telephone network(PSTN). The SIP protocol itself is modeled on the three-way handshake method (see Figure. (1).

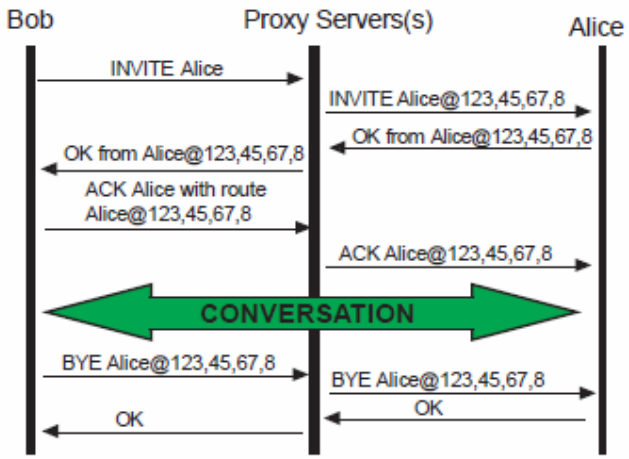


Figure.1. Session initiation protocol.

Consider the setup when a proxy server is used to mediate between endpoints. During the setup process, communication details are negotiated between the endpoints using Session Description Protocol (SDP) which contains fields for the codec used, caller's name, etc. If Bob wishes to place a call to Alice he sends an INVITE request to the proxy server containing SDP information for the session, which is then forwarded to Alice's client by Bob's proxy, possibly via her proxy server. Eventually, assuming Alice wants to talk to Bob, she will send an "OK" message back containing her call preferences in SDP format. Then Bob will respond with an "ACK". SIP provides for the ACK to contain SDP instead of the INVITE, so that an INVITE may be seen without protocol specific information. After the "ACK" is received, the conversation may commence along the RTP / RTCP ports previously agreed upon, SIP presents several challenges for firewalls and NAT.

2.2. Internet Protocol Security

IPsec is a suite of protocols for securing Internet Protocol (IP) communications by encrypting and/or authenticating each IP packet in a data stream [4]. IPsec also includes protocols for cryptographic key establishment. IPsec is implemented by a set of cryptographic protocols for securing packet flows and Internet key exchange. IPsec was intended to provide either transport mode: end-to-end security of packet traffic in which the end-point computers do the security processing, or tunnel mode: portal-to-portal communication security in which security of packet traffic is provided to several machines (even to whole LANs) by a single node. IPsec can be used to create Virtual Private Networks (VPN) in either mode, and this is the dominant use. The security implications are quite different between these two operational modes. IPsec was introduced to provide security services such as: Encrypting traffic (So it can not be read in its transmission) Integrity validation (Ensuring traffic has not been modified along its path).Authenticating the Peers (Both ends are sure they are communicating with a trusted entity the traffic is intended for)[8]. Two main factors affect voice traffic when IPsec is used. The first one is the increased packet size because of the headers added to the original IP packet, namely the ESP header for confidentiality and the new IP header for the tunnel. The second one is the time required to encrypt payload and headers and the construction of the new ones. this section report the influence of such factors on voice traffic. Realistic estimates of such factors can be determined only through a careful experimental analysis, as most of the parameters involved such as traffic shape, buffering delay and queuing delay depend on real traffic condition. Figure. 2 illustrate the format of voice packets with various protocols for a 40 bytes payload, a typical packet length for voice traffic. The figure shows how packet format and size change with and without IPsec and for various combinations of cryptographic algorithms. The overall minimum size is obtained when compressed RTP (cRTP) is adopted (second bar from the bottom) in which case the header size is only 20% of the payload size, yielding a 45 bytes long packet, while a regular IP packet is 80 bytes long. As it can be seen, the use of IPsec dramatically increases the size of the packets, which reaches 120 and 130 bytes depending on the cryptographic services requested. As a consequence, the ratio between the actual payload and the total packet length decreases, indicating an increase in “wasted” bandwidth.

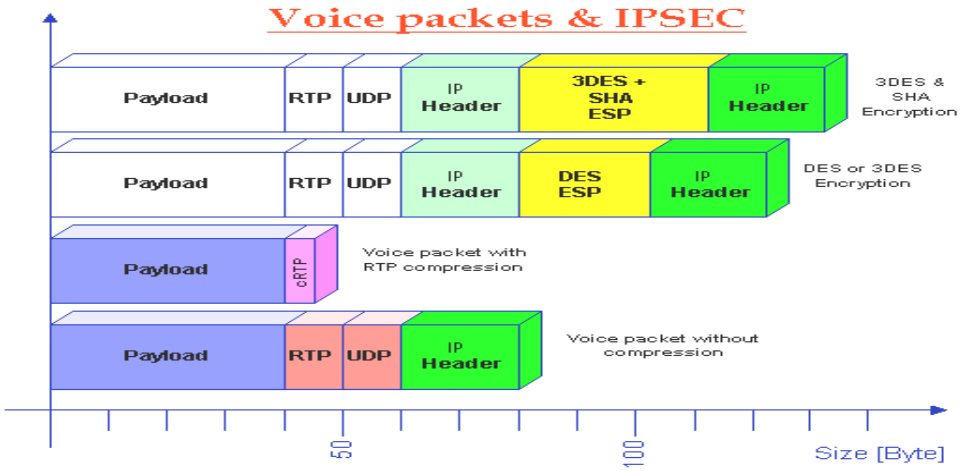


Figure.2. Voice packet format [2]

During the encryption process, a router performs some operations, namely packet encryption and new headers construction (ESP + IP tunnel), that influence the CPU utilization and introduce a further delay. the crypto-engine is a serious bottleneck in the transmission of real-time traffic in IPsec.The main reason, however, is not the low efficiency of the encryption process but the difficulty to control packet access to the crypto-engine.

3. VoIP Security Requirements and Quality Metrics

As every VoIP network is essentially an IP network, VoIP network and terminals face the same security threats inherent with any IP network. For example, the media (RTP) packets of an ongoing VoIP call on a LAN could be easily picked up and recorded by a simple packet sniffer. In naive terms, VoIP calls need to be at least as secure as circuit-switched calls with respect to anonymity and privacy. With added security threats because of the open nature of the underlying IP network, the functional VoIP security requirements could be stated as:

(1) Protection of privacy of the call conversation.
(2) Authentication of call end entities.
(3) Protection from misuse of network resources, or in other words, access control by the service provider.
(4) Ensuring correct billing by the service provider, and protecting billing information from unauthorized access.
(5) Protection of caller behavior or statistical information from unauthorized access.
(6) Protection of network servers and terminals from well known threats, such as ‘denial of service’ and ‘man in the middle attacks’.

In essence both the media stream and the signaling stream of a VoIP call must be protected from unauthorized access. Most of the security requirements highlighted above are not specific to VoIP alone, but are general requirements of any IP network wishing to protect the interests of its end entities. However, unless the underlying network is secure, either the VoIP call is not secure, or it has to have its own security features implemented .

An amount of research has been carried out into the issue of providing security protocols for specific applications on IP networks (e.g. Secure Socket Layer (SSL), which is primarily used in eCommerce applications , Pretty Good Privacy (PGP), which is primarily used for email)
And IPsec, which is primarily used in VPN(Virtual Private Network), These already developed security protocols could be employed for VoIP as well. However, the characteristics and resource requirements of securing voice traffic (and indeed multimedia traffic in general) are distinctly different from traditional IP traffic.

In general, the available options for VoIP security are:

(7) Integrating the core security mechanisms of authentication and encryption, into the VoIP protocols itself.
(1) Use existing application layer security protocols (e.g. PGP), or similar for providing security services to VoIP signaling and media streams.
(2) Generic transport layer security protocols, like SSL/TLS could be employed.
(3) More flexible and scalable alternative is to carry VoIP over ‘secured’ networks, which mean the using of security services built into the network layer (IPSec).

The quality of service (QoS) is one of the main characteristics that are important for people Using IP telephony (VOIP). These includes several Metrics such as delay, jitter ,packet loss and call setup time. Since the voice is transferred in real-time it is important to control jitter and delays in order to guarantee QoS for IP telephony networks

- **Latency**: As a delay-sensitive application, voice cannot tolerate too much delay. Latency Is the average time it takes for a packet to travel from its source to its destination. The Maximum amount of latency that a voice call can tolerate one way is 200 milliseconds (150 Milliseconds is preferred)[6] . If there is too much traffic on the line, or if a voice packet gets stuck behind a bunch of data packets (such as an email attachment), the voice packet will be delayed to the point that the quality of the call is compromised.

- **Jitter**: In order for voice to be intelligible, consecutive voice packets must arrive at regular Intervals. Jitter describes the degree of variability in packet arrivals, which can be caused By bursts of data traffic or just too much traffic on the line. Jitter is the delay variance from point-to-point. Voice packets can tolerate only about 75 milliseconds (50 milliseconds is preferred) of jitter delay .

- **Packet loss**: Packet loss due to congestion is the losing of packets along the data path, which Severely degrades the voice quality. Packet loss occurs frequently in data networks, but many applications are designed to provide reliable delivery using network protocols that request a Retransmission of lost packets (e.g. TCP). Dropped voice packets, on the other hand, are Discarded, not retransmitted. Voice traffic can tolerate less than a 10 percent loss of packets before callers feel perceivable gaps in conversation.

4. Performance Analysis

4.1 Simulation Environment

The simulation experiment is carried out using OPNET simulator under Windows XP as a platform, the OPNET instructions can be used to define the topology structure of the network. Fig. 3 shows the reference structure of the SIP–VoIP network model built The shown network topology consists of three different sites, ‘site1’ to ‘site3’, every pair of which is interconnected by an IP router. The SIP proxy server with a co-located location server acts as the functional core. The first site act as caller nodes ,the second site act as SIP proxy server and the third site act as callee nodes. The main components that build up the network model are nodes(caller or callee), sip proxy server,IP routers and IP cloud which represent the Internet.

4.2 Simulation Results And Analysis

This section reports the results obtained to examine the impact of IPsec on the quality of transmitting voice over communication links for different simulation times using OPNET simulator. The measuring criteria’s used to evaluate the mentioned protocol are end to end delay, packet delay variation (jitter) , packet loss and call setup time.

The maximum acceptable delay in packet delivery for optimal voice quality is 150ms which can be extended up to 200ms in case of encrypted communication. Figure. 5 shows the measured end to end delay for plain IP, IPsec with authentication only, IPsec with encryption

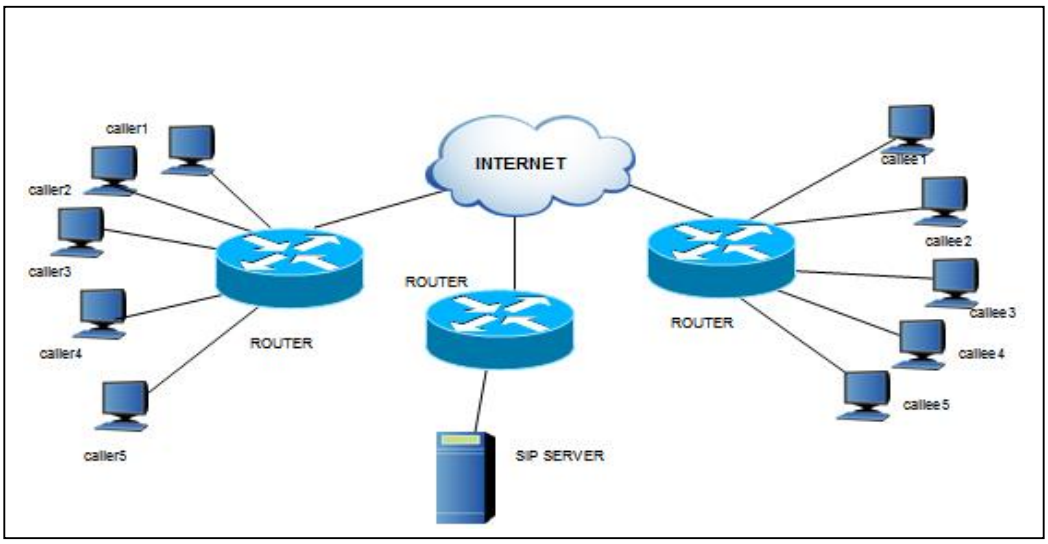


Figure.3. The structure of SIP_VOIP network model

only and IPsec with both encryption and authentication for different simulation times. It could be noticed that the end to end delay time is approach to 115 ms for plain IP, 140 ms for IPsec with authentication only, 175 ms for IPsec with encryption only and 200 ms for IPsec with both encryption and authentication.

The second parameters measured is packet delay variation(jitter) the maximum acceptable packet delay variation(jitter) for optimal voice quality is (40:75)ms Figure.6 shows the measured packet delay variation(jitter) for plain IP, IPsec with authentication only, IPsec with encryption only and IPsec with both encryption and authentication for different simulation times. It could be noticed that the packet delay variation(jitter) is approach to 40 ms for plain IP, 50 ms for IPsec with authentication only, 75 ms for IPsec with encryption only and 85 ms for IPsec with both encryption and authentication.

The third parameters measured is packet loss the maximum acceptable packet-loss is less than 10%, Figure.7 graphs the measured packet loss for plain IP, IPsec with authentication only, IPsec with encryption only and IPsec with both encryption and authentication for different simulation times. It could be noticed that the packet loss is approach to 4% for plain IP, 6% for IPsec with authentication only, 8% for IPsec with encryption only and 10% for IPsec with both encryption and authentication.

The last parameters measured is call setup time , The maximum acceptable call setup time is less than 32sec. It could be noticed that from figure. 4 the call setup time approach to 1 sec for plain IP, 2sec for IPsec with authentication only, 7.5 sec for IPsec with encryption only and 10sec for IPsec with both encryption and authentication. Figure.4 The call setup time for plain IP, IPsec with authentication only, IPsec with encryption only and IPsec with both encryption and authentication for different simulation times. (a) 1 hour simulation (b) 5 hour simulation (c) 24 hour simulation. Figure(5) The End to End delay for plain IP, IPsec with authentication only, IPsec with encryption only and IPsec with both encryption and authentication for a) 1 hour simulation, b) 5 hours simulation c) 24 hour simulation. Figure(6) The packet delay variation(Jitter) for plain IP, IPsec with authentication only, IPsec with

encryption only and IPsec with both encryption and authentication for a) 1 hour simulation, b) 5 hours simulation c) 24 hour simulation. Figure(7) The packet loss for plain IP, IPsec with authentication only, IPsec with encryption only and IPsec with both encryption and authentication for a) 1 hour simulation, b) 5 hours simulation c) 24 hour simulation.

5. Conclusion

This paper has described a simulation model of an IPSec secured SIP based VoIP network. Performance analysis of the different configurations of IPSec for VoIP networks has been evaluated through a series of experiments on OPNET simulator. the impact analysis of employing IPsec encryption and authentication services for VOIP signaling and media streams shows that between encryption and authentication, encryption is the more expensive operation in End to End Delay, Packet Delay Variation(Jitter),Packet Loss and Call Setup Time. When both encryption and authentication services are employed a dramatically increase in call setup times, an increase of around 170% in the media stream delay and around 200 % in jitter value, the SIP call setup time and media stream delay seem to increase exponentially with increase in the network call density. This increase is mainly due to the packets getting queued up at routers waiting to be processed by the IPsec security engine. Experiments with other types of real-time traffic results presented in this paper can be generalized to all real-time traffic are a part of future work .

6. References

[1] Ashraf D. Elbayoumy and Simon J. Shepherd, " QoS control using an End-Point CPU capability detector in a secure VoIP system", In Proceedings of the 10th IEEE Symposium on Computers and Communications, 2005.

[2] R. Barbieri, D. Bruschi and E. Rosti, "Voice over IPsec: Analysis and Solutions", Computer Security Applications Conference, 2002. Proceedings. 18th Annual, 9-13 Dec. 2002.

[3] M. Leelanivas, "RTP Header Compression", Cisco Systems, 1997.

[4] D. Nguyen and J. Lequang, "cRTP Performance Enhancement", Cisco System [ENG 102721],2002.

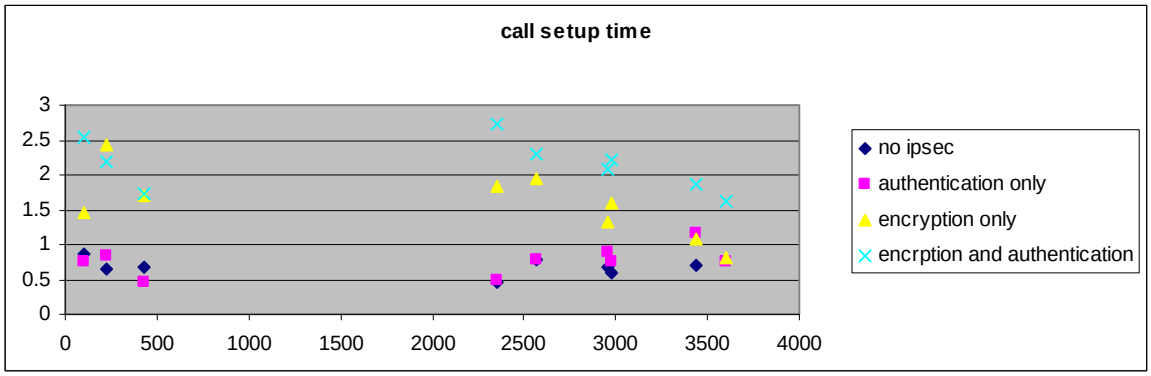
[5] Seema Ansari and Arshi Khan, " Voice over Internet Protocol Security Problems in Wireless Environment", PAF-KIET Journal of Engineering and Sciences Volume 01, Number 02, July-December 2007.

[6] G.114: One-way Transmission Time, ITU-T Recommendation, G Series, 2000.

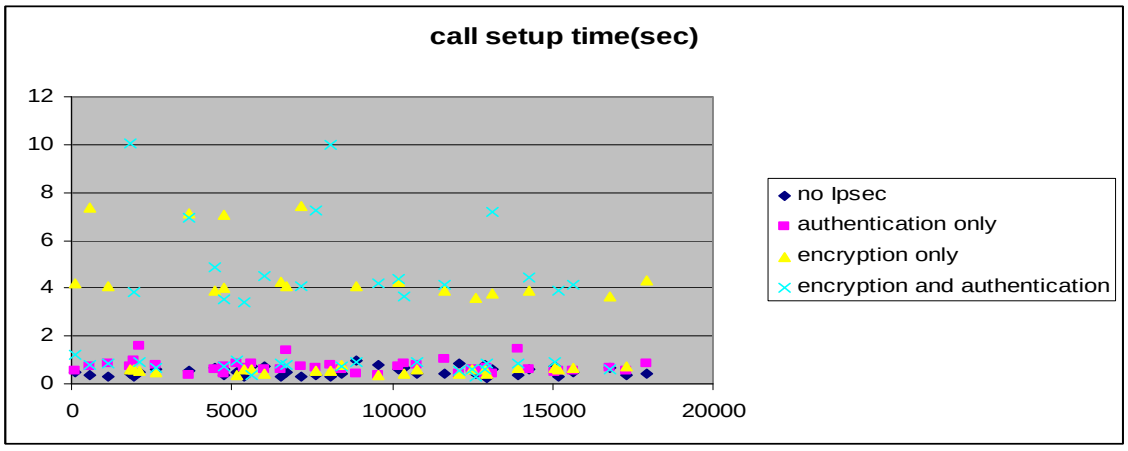
[7] J. Casteel. Sound Choices for VoIP Security, 2005. Retrieved October 20, 2008 from

[8] http://www.ebcvg.com/pdf/dl/sound_choices_voip_security.pdf

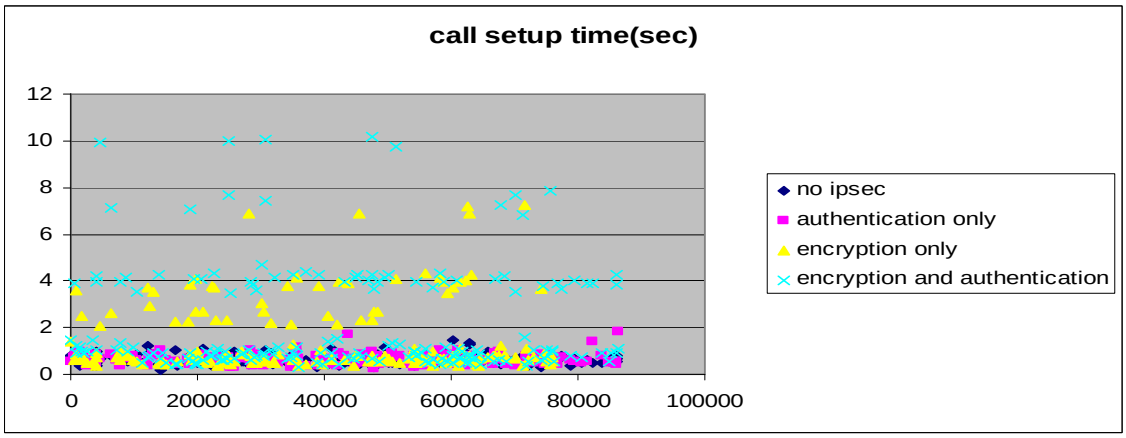
[9] Jirka Klaue and Andreas Hess, "On the Impact of IPsec on Interactive Communications", In Proceedings of the 19th International distributed processing Symposium, 2005.



(a)

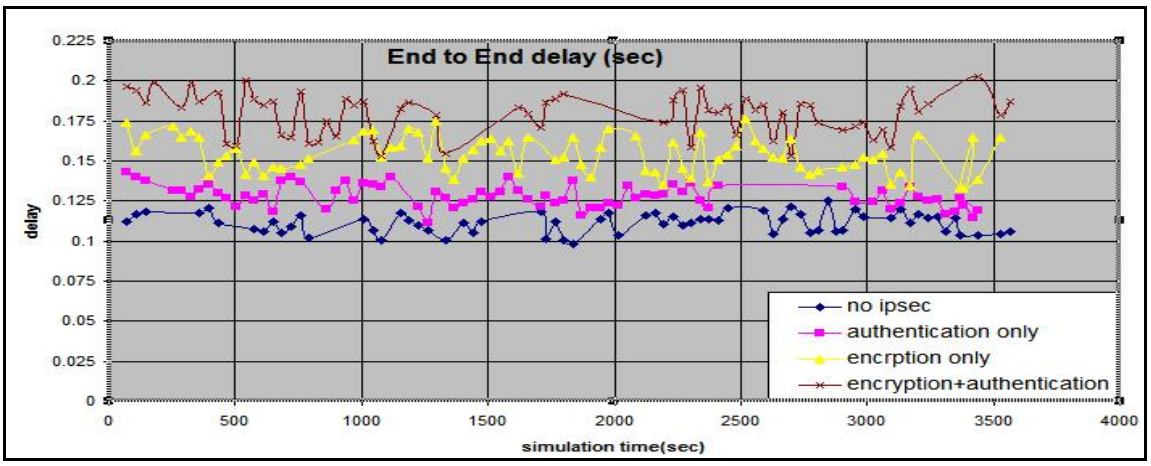


(b)

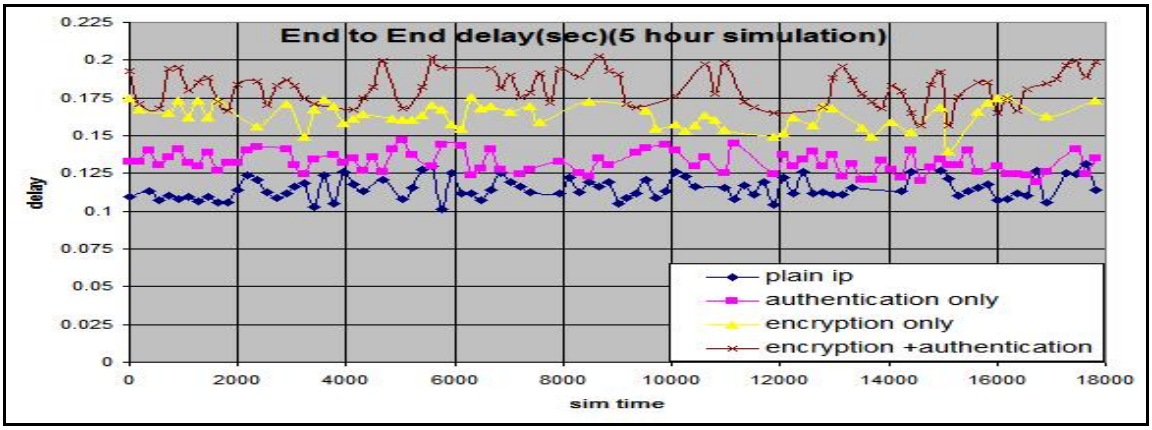


(c)

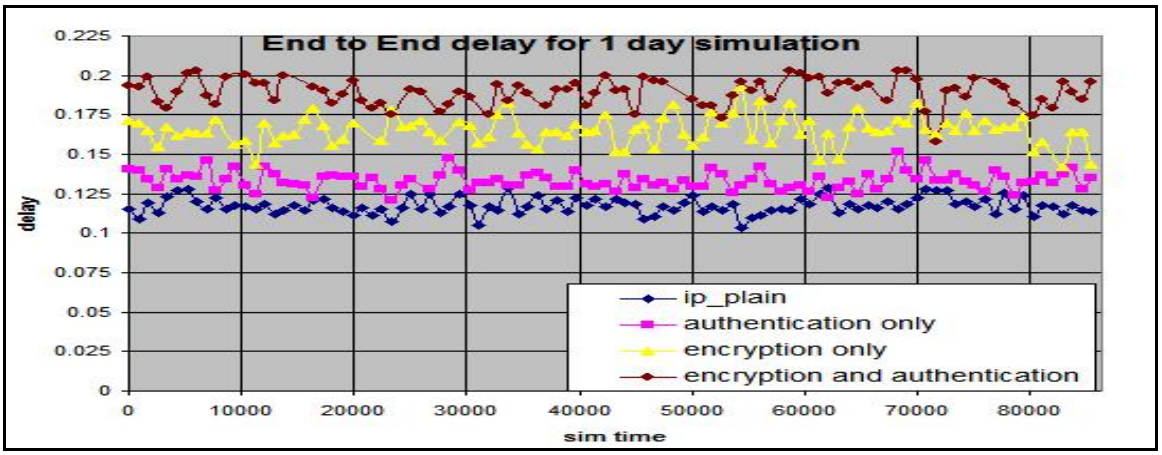
Figure.4. Call setup time



(a)

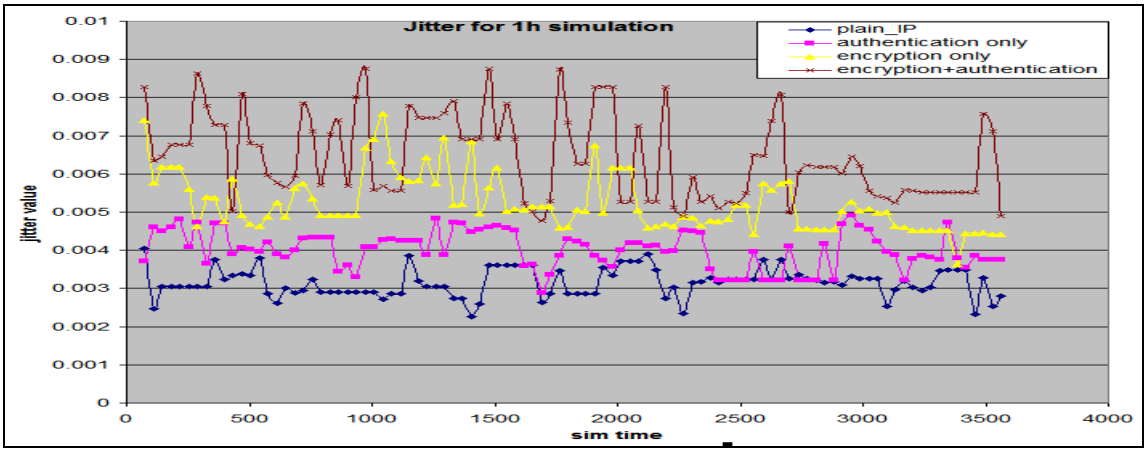


(b)

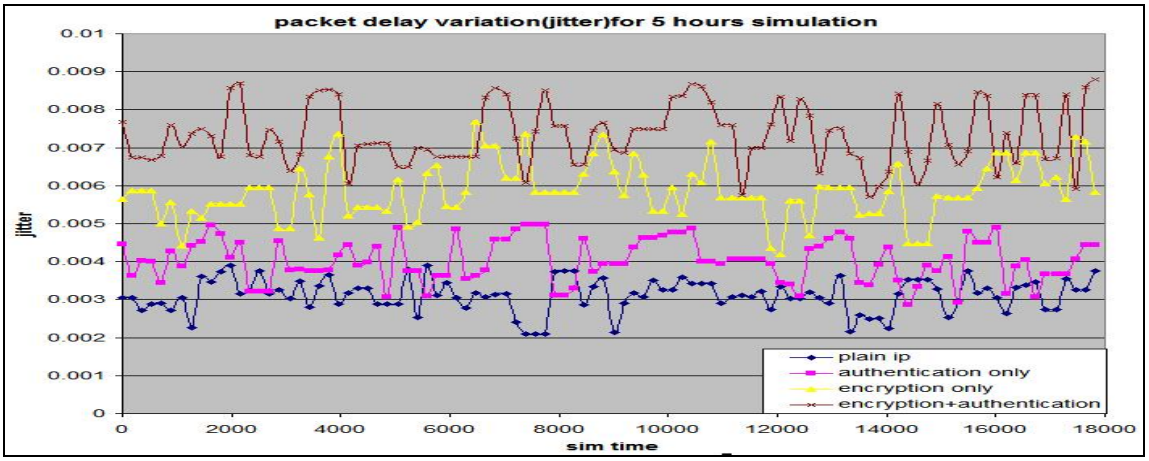


(c)

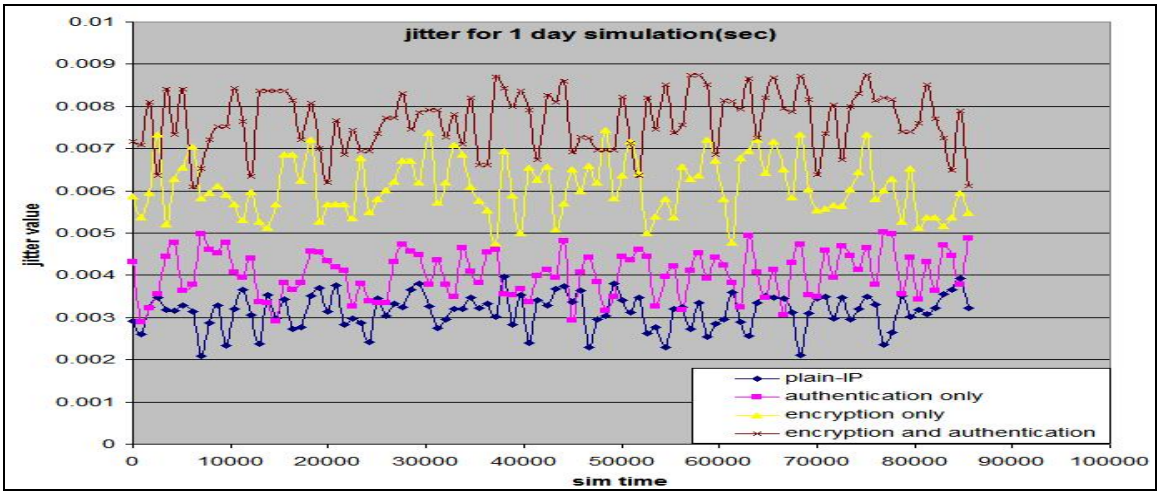
Figure.5. End to end delay



(a)

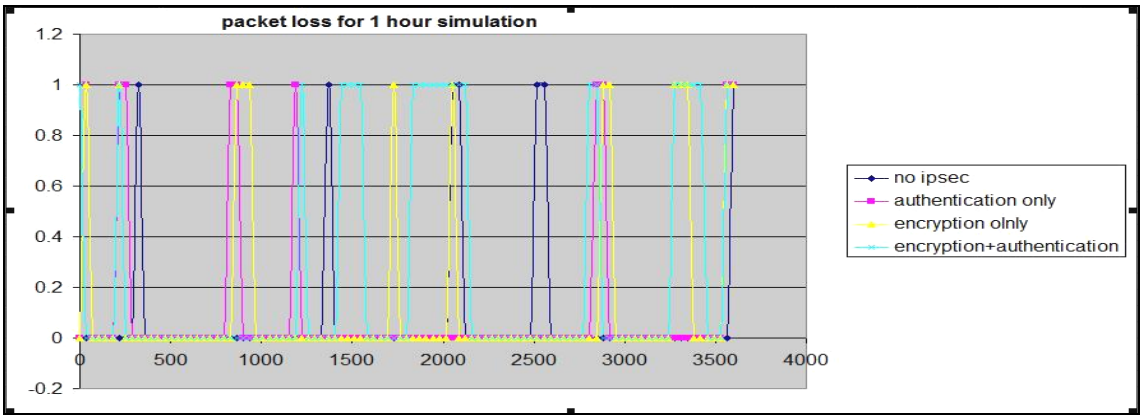


(b)

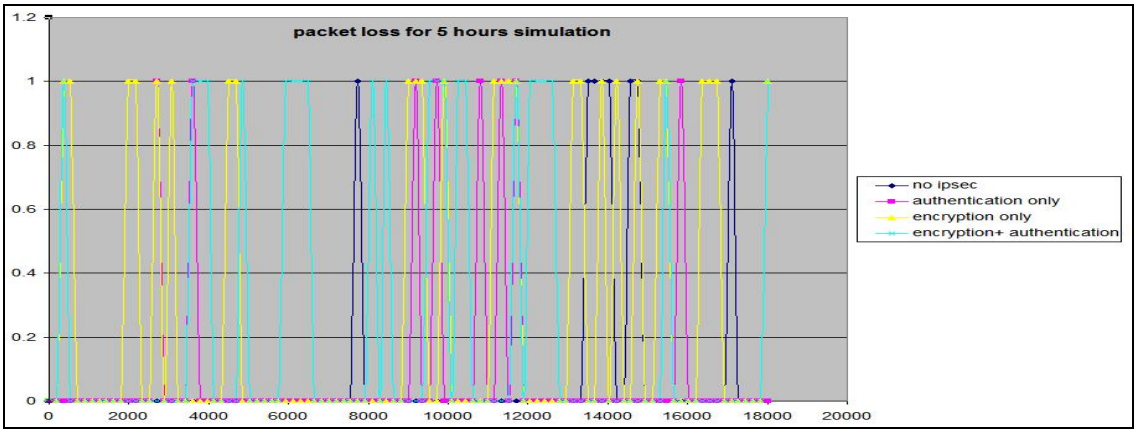


(c)

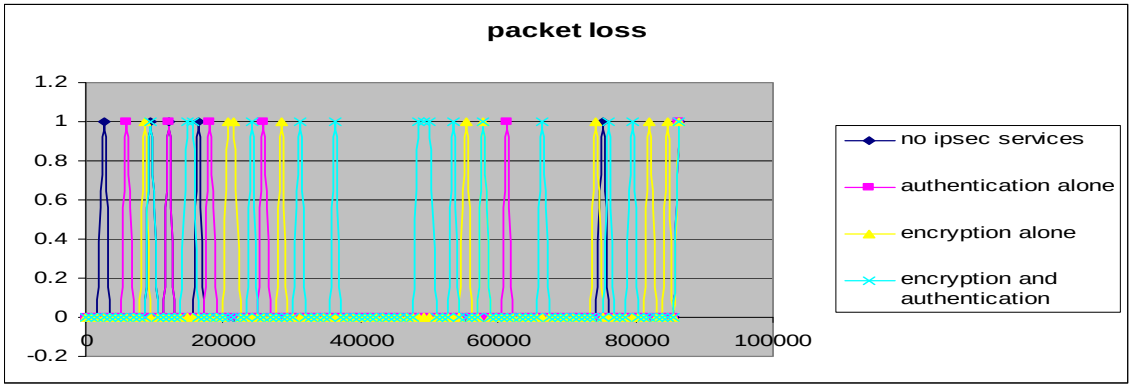
Figure.6. Packet delay variation(Jitter)



(a)



(b)



(c)

Figure.7. Packet loss