Menemui Matematik (Discovering Mathematics)
Vol. 34, No. 2: 57– 65 (2012)

Performance Analysis of High-Definition Video Call over Secure Real Transport Protocol (SRTP)

Nor Rahimah Ismail and Saadiah Yahya
Faculty of Computer Science & Mathematic,
University Technology MARA, Shah Alam, Selangor
imahismail@gmail.com, saadiah@tmsk.uitm.edu.my

ABSTRACT

The deployment of high-definition (HD) video requires the employment of security for ensuring availability, confidentiality and reliability communication. The Session Initiation Protocol (SIP) is considered as principal signalling protocol for handling real time data, voice and video over the Internet. SIP like other Internet protocol is vulnerable to Internet attacks. Meanwhile the Secure Real-Time Transport (SRTP) is an extension of RTP protocol which is used to protect the media flows such as voice and video. SRTP was designed to provide the protection and confidentiality of RTP packet. Nowadays most of the software developers have lack of concern about the security features to be added into their video call applications may be due to their interest in the performance of video quality delivery. Thus the objective of this research is to measure and analyze the performance of HD video call on two different channels which are the secure and the non-secure channel. The experiments are conducted on wired and wireless environment. In this research, three indicators were used namely: jitter, MOS score and R-Factor. The experiments were conducted using *CounterPath Bria Professional softphones* on two clients together with a SIP server on BSD platform to enable and disable SRTP for secure and non-secure channel configuration. A simulation of high-definition video over IP environment has been created and the results taken are measured. The findings reveal that the performance for the non secure channel gives better result compared to the secure channel. However, the HD video call application is acceptable to be used in the secure channel even though the secure channel has higher jitter than in the non-secure channel. This is evidenced when the security features were added into a secure channel, the process of encryption and decryption has to be performed. For the future research, this study can be extended over WANs or other kind of network.

Keywords: High-Definition, Secure Real Transport Protocol, Performance, Environment

INTRODUCTION

After so many decades the telephone technology has developed such a valuable technique that helps people in making multimedia telecommunicating. Many organizations have started to realize that they can save money by moving video and voice traffic over IP. Thus, many broadcasters switched to IP for video transmission. HD video traffic is also deployed over IP for video call application. Previously, HD video has already being used in the television system for the higher resolution. HD video call over the internet is then deployed as it promised good quality and resolution. Currently, HD video is one of the newest technology available and it already over takes the standard-definition (SD) video when it became increasingly used by users due to the pleasure of high resolution video with high quality. Video calls over internet is feasible since video and voice can be integrated which can reduce cost and management effort. Thus, many developers have deployed the applications with lack of considerations on the security issues. Similarly, high definition implementation in video call ensures quality delivery of video images. However, it did not offer security features like the voice over IP had. Correspondingly, users do not aware on the importance of delivering video call over security layer and the benefit of using this application. It will be good when many people realize and become conscious of using video call over IP technology. They need to also ascertain the security of this video over IP applications. Therefore, the development of a secure channel using secure real-time protocol is necessary.

The objectives of this research is to determine and analyze the performance of HD video call on two different channels which are secure and non-secure channel and to deploy HD video call over a secure network environment. 802.11n wireless is used as the transmission technology and the CounterPath Bria Professional software that act as video call over IP soft phone is installed at two laptops. This Bria Professional software can support SIP signaling and SRTP media encryption. Meanwhile, for the analysis purposes, CommView analyzer for video call over IP is used. The video display resolution for the image is set to 720 pixels.

HD Video Overview

The HD standard has been developed by Advanced Television Systems Committee (ATSC) and was adopted by the Federal Communications Commission (FCC) in year 1996 [1]. Because of its higher resolution and richer color makes the HD video more popular to the users than standard-definition video. Normally the common used for display resolution of the HD video are 1280x720p and 1920 x 1080 pixels (1080i/1080p) [2]. According to [3] in 720p resolution, the number '720' stands for the number of horizontal scan lines of display and the letter 'p' stands for progressive scan. The widescreen aspect ratio of this resolution is 16:9. It contains a vertical resolution of 720 pixels and a horizontal resolution of 1280 pixels, for a total of 921,600 square format pixels. The 1280 × 720 format is always progressive scan, where the entire frame is scanned sequentially from top to bottom in horizontal direction. The commonly used frame rates for this resolution are 23.976, 24, 25, 29.97, 30, 50, 59.94, and 60 frames per sec. Meanwhile for the 1080i and 1080p the HD resolution has 1080 horizontal scan lines of display with interlace scan and progressive scan correspondingly. For the interlace scan, the display is divided into two fields that is odd horizontal lines and even horizontal lines and it is processed one after another [3]. Therefore, to transmit the raw HD video without any compression, the requirement bandwidth needs should be around 4 Gbps.

SRTP: Securing Media Flow

Secure RTP is defined a profile for RTP that enable message authentication, add confidentiality and replay protection to that protocol. A number of studies showed that SRTP is a good option to secure a media session [4]. Furthermore, the security is added at the application layer. SRTP is profiles that secure the RTP media transport protocol specified by the proposed standard in the RFC 3711 [5]. Even though the encryption is mandatory, this protocol is thought to offer confidentiality, integrity and authentication to the packets of an RTP flow. Figure 1 show the security is applied on the packet. By using SRTP protocol to protect the RTP media, had promised to overcome some security issues such as confidentiality, authentication and integrity.

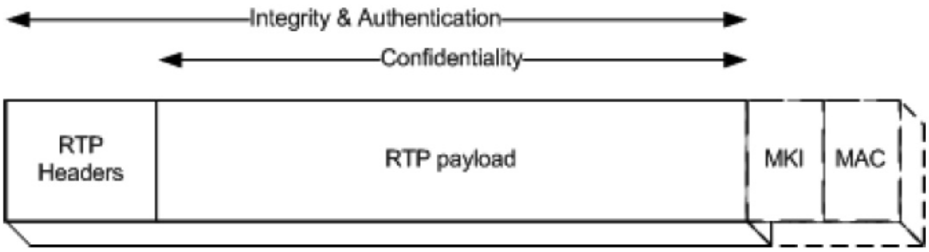

Figure 1: SRTP schema

Issues in High-Definition Video Call Implementation

There are several issues need to be considered in HD video call implementation which can be seen as the limitation and the sequences of using this application. There are 3 main causes affecting the video and voice quality namely: jitter, MOS score, and R-factor, which are discussed in the following subsections.

Jitter

Jitter refers to a non-uniform packet delays, where the information is transferred from source to destinations via small messages called packets. The packets experience certain delays to reach their required destinations. The variation in these delays is known as jitter and it favorably affect quality of the service provided. It significantly degrades quality due to packet loss but can be managed via jitter buffers. At the sending side, packets are sent in a continuous stream with the packets spaced evenly apart. Generally, jitter is caused by the congestion in the IP network. It can occur either at the router interfaces or in a provider or a carrier network if the circuit has not been provisioned correctly [6]. The maximum acceptable level of packet jitter is 30ms for video applications and 10-20ms for voice application [7].

MOS Score

Mean Opinion Score is one of the methods used to quantify the voice and video quality at the destination end of a communication link. MOS is based on averaged ratings by a sample of listeners in expressing voice and video quality whether it is good or bad. MOS can be tested using a simulation model, human perception and automated system [8] [9]. MOS shows a numerical sign of the quality of the media received after being transmitted and expresses it in one number from1 to 5 as shown in Table 1 [10], [11].

Table 1: Mean Opinion Score (MOS) Ratings

	Mean Opinion Score (MOS) Ratings
Excellent	5 (Perfect - Like face-to-face conversation or radio reception)
Good	4 (Fair - Imperfections can be perceived, but sound still clear. This is (supposedly) the range for cell phones)
Fair	3 (Annoying)
Poor	2 (Very annoying. Nearly impossible to communicate)
Bad	1 (Impossible to communicate)

R-Factor

E-Model has been introduced by the recommendation of ITU-T G.107 as an alternative rating scale to MOS. The scale rating for using R-Factor over MOS is from 1 to 120 encompassing both the narrow and wideband Codec. R-Factor quantify individual impairments such as jitter, delays, echo, packet loss, codec type, noise and equipment factors and gives a single number [12][13]. R-Factor is more consistent than MOS and its produce more consistent scores.

RELATED WORKS

The implementation of SRTP is able to improve the VoIP quality. Mohd Nazri Ismail [14] has conducted experimentation and analysis of a one to one and multi conference video call communication using 5 selected codecs which are G.711, GSM, G.726, SPEEX and iLBC. The results confirmed that performance was affected when security features were enabled on all the codecs. Andre L. Alexander *et al.* [15] defines the performance of SRTP and its impact on VoIP in LAN environment with and without SRTP. The researchers then measure jitter, delta and VoIP throughput. With VoIP, there is no operation degradation when SRTP is implemented. However, SRTP adds negligible overhead to VoIP and does not affect the VoIP qualities in terms of packet inter arrival time and jitter. Meanwhile, Sureshkumar and Rudra Dutta [16] have evaluated key performance parameters namely: call setup time, mean number of calls, memory utilization and queue size. They concluded that, the additional security implemented on calls can impact performance of those parameters. T. Adomkus and E.Kalvatis [17] also confirmed that SRTP can deprive VoIP services in terms of packet delay, throughput and also utilization, however the result for delay did not show a big impact. They also proposed that VoIP over SRTP is necessity to be used for voice encryption.

EXPERIMENT

The main objective of the study is to determine and analyse the performance of HD video call on two different channels which are secure and non- secure channel. The study involved a controlled LAN which consists of three laptops. A SIP server running on BSD platform is using Acer Ferrari 1000 AMD Turion™ 64 X2 Dual-Core Mobile Technology TL-62 with 120 GB Hard disk and 3 GB of RAM computer. The laptops were installed with Windows XP operating system and VMware version 6.5 were used to setup the SIP server. For the clients two laptops were used and installed with; Acer Aspire 4736 Intel® Core™ Duo Processor T6600 2.2 GHz with 132 GB Hard disk and 3 GB of RAM and HP Elite Book Intel® Core 2 Duo Processor 2.4 GHz with 200 GB Hard disk and 3 GB of RAM. Both laptops were running on Windows XP as a platform. The OpenSER SIP server was used to register the clients and to set up video calls over IP between the clients. In the meantime, the QoS parameters were measured using CommView version 6.0 analyzer. For the softphone, Bria Counterpath Professional version 2.4.3 was used and installed at two clients to enable the HD video call. In addition, G.711 is used as an audio codec. Since this research focus on HD video, Logitech HD Webcam C270 was installed at both laptops. This webcam is solidly built in and offers 720p HD video quality. Meanwhile, to setup the wireless environment, Linksys Wireless-N Home Router WRT120N is used to enable the communications between nodes. At the same time, Linksys Dual-Band Wireless-N USB Adapter WUSB600N is installed at clients since the simulations were done in wireless-N. In this research each phase need to be done repeatedly for each video calls for ten times in six days; three days for wired and three days for wireless. The experimental setup involved three nodes; one for SIP server and two for clients. This simulation has been measured and analyzed by predefined QoS namely jitter, MOS score and R-Factor.

RESULTS AND FINDINGS

The experiments conducted were able to validate performance comparison between wired and wireless HD video call. This analysis was based on video and voice over IP network. The experimentation involved data gathering which then has been tabulated and analyzed. The findings were based on the indicated three parameters on two different environments; wired and wireless.

Jitter

As shown in Table 2 and Figure 2 below, the readings for average jitter in secure channel is higher than in non-secure channel. This result shows that performance of HD video call for non-secure channel is better than in secure channel. Even though jitter for secure channel is higher than non-secure channel, it does not mean that the call cannot be done, but the call is acceptable for both channels to be made. When the security features were added in this video call application, the process of encryption and authentication were involved. This has resulted in slow response and packet delay. However the jitter value is almost negligible and can still be acceptable since the threshold level of packet jitter is 30ms for video applications and 10-20ms for voice application [18].

Table 2: Average readings for Jitter in wired

	Non Secure	Secure
Day 1	2.44ms	2.944ms
Day 2	2.453ms	2.574ms
Day 3	2.2965ms	2.707ms

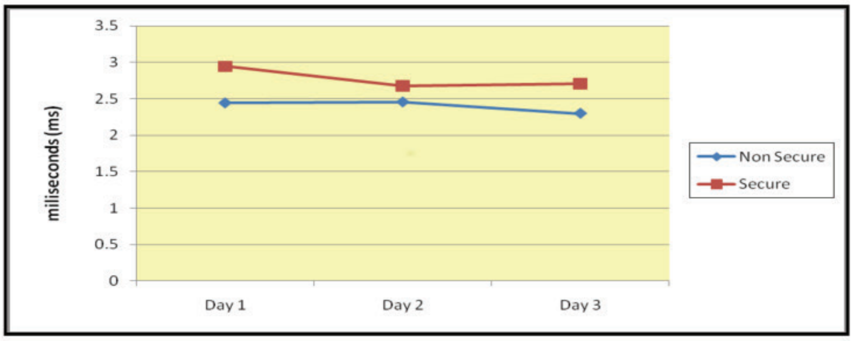

Figure 2: Average Readings for Jitter on wireless

Meanwhile readings recorded in Table 3 and Figure 3 shows that jitter for video call in secure channel is significantly higher than the non-secure channel. Even though jitter for secure channel is higher than non secure channel, it does not mean that the call cannot be done. However, as explained earlier, the call is acceptable for both channel since the threshold level of packet jitter for video applications is 30ms. Additionally, the result for jitter in wireless shows higher value compared to the wired one. Interference and congestion that is normally experience by wireless networks can limit video transmission performance.

Table 3: Average Readings for Jitter on wireless

	Non Secure	Secure
Day 1	11.7405ms	13.307ms
Day 2	13.0885ms	13.638ms
Day 3	10.9655ms	11.4275ms

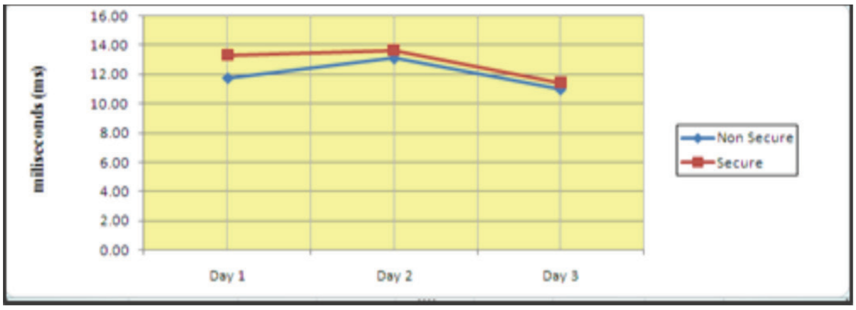

Figure 3: Average Readings for Jitter on wireless

MOS Score

Table 4 and Figure 4 illustrate the average of MOS score reading both channels. The results show the reading for non-secure channel is more reliable for the transmission with score 4.4. Instead, result for security channel showed the average reading is 4.39.

Table 4: Average Readings for MOS Score on wired

	Non Secure	Secure
Day 1	4.4	4.38
Day 2	4.4	4.39
Day 3	4.4	4.39

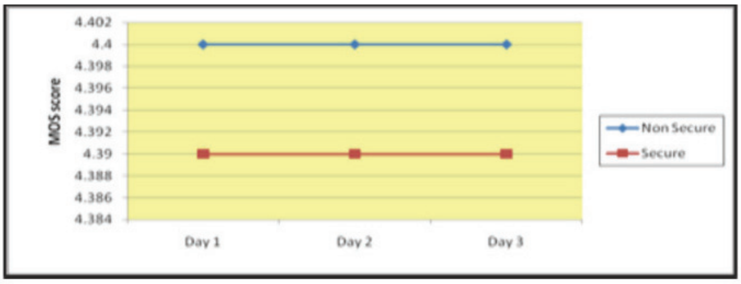

Figure 4: Average Readings for MOS Score in wired

Table 5 and Figure 5 illustrate the average of MOS score taken for both channels during the experimentation. The result shows that the MOS score for secure channel is significantly less than MOS score for the non-secure channel. It means that, the call quality is much better in the non-secure channel. However, secure channel can still be considered since MOS score value is still lies in a good quality range as shown in Table 1.

Table 5:Average Readings for MOS Score in wireless

	Non Secure	Secure
Day 1	4.39	4.365
Day 2	4.4	4.39
Day 3	4.4	4.39

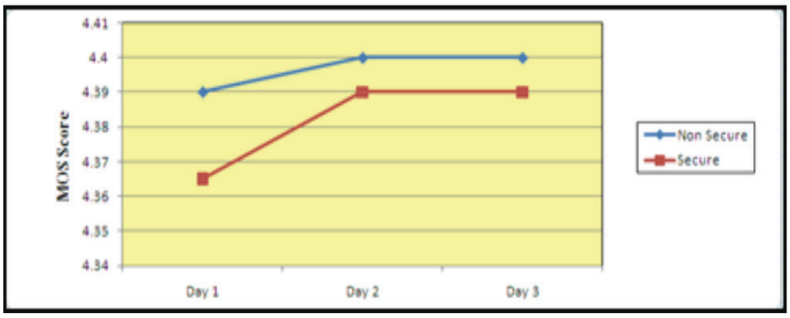

Figure 5: Average Readings for MOS Score on Wireless

R-Factor

Meanwhile Table 6 and Figure 6 illustrate the average number of R-Factor for both channels. For day one, result for non-secure channel is 93.255 and secure channel is 93.155. For day two, results for non-secure channel is 93.2 and secure channel is 93.155. Finally, day 3 results show that, the non-secure channel is 93.165 and the secure channel is 93.065. The findings from the analysis conclude that, the performance for the non-secure channel is better than the secure channel. Even though the value for secure is lower than the non-secure channel, the video call is still acceptable as been discussed in the MOS score simulation section above.

Table 6: Average Readings for R-Factor on wired

	Non Secure	Secure
Day 1	93.255	93.155
Day 2	93.2	93.155
Day 3	93.165	93.065

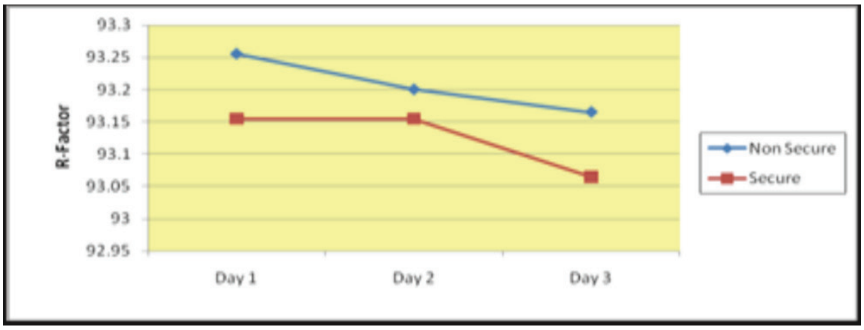

Figure 6: Average Readings for R-Factor on wired

In the wired connection, readings for both channels are consistently scored 93 but for wireless we found that the readings are slightly varied. 10 times readings are recorded and the average for both channels is calculated as shown in Table 7. Correspondingly, Figure 7 shows that SIP video calls in non-secure channels give better result compared to the secure channel.

Table 7: Average Readings for R-Factor on wireless

	Non Secure	Secure
Day 1	93.155	92.91
Day 2	93.2	93.07
Day 3	92.975	92.77

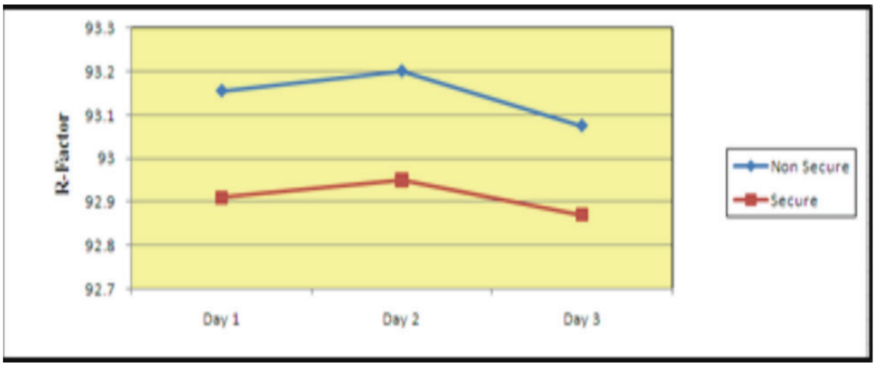

Figure 7: Average Readings for R-Factor on Wireless

CONCLUSION

The study focuses on the analysis of the HD video calls over secure and non-secure channel LAN network using jitter, MOS score and R-Factor as performance indicators. The findings from the analysis show that the performance in the non-secure channel is better with wired and wireless communication. However, the result for the jitter in secure channel give higher value compared to the non-secure channel. The reason for this is, when the security features were added in the secure channel, certain provision such as the process of encryption and decryption is taken place in the network. Though the HD calls in secure channel do not give better performance readings as in the non-secure channel, the calls can still be acceptable. This is proven when the MOS score and R-Factor also give good results for both channels. Thus, HD video call is capable to transfer calls on secure channel LAN networks.

REFERENCES

Wu, Y., S. Hirakawa, U. H. Reimes, and J. Whitaker, *"Overview* of digital television development worldwide", Proceedings of the IEEE, January 2006.

Jack, K. "Video Demystifed a Handbook for the Digital Engineer", 4th edition, Newnes, USA, 2005.

Parhi, K. K. and T. Nishitani (eds), "Digital Signal Processing for Multimedia Systems" Marcel Dekker,USA, 1999.

Orrblad, Joachim, "Alternative to MIKEY/SRTP to secure VoIP", Royal Institute of Technology (KTH), Telecommunications System Laboratory, 2005.

Baugher, M. and others, "The Secure Real-time Transport Protocol (SRTP)", Internet Engineering Task Force, RFC 3711, 2004.

Cisco System," Understanding Jitter in Packet Voice Network",2007.

André Leroux, Bruno Giguère, M "SONET/SDH vs Ethernet: Migration and Testing Issues", 2006.

Moura N.T.; Vianna B.A.; Albuquergue C.V.N; Rebello V.E.F & Boeres C. "MOS-Based Rate Adaption for VoIP Sources". IEEE International Conference on Communication, 2007.

Masuda M. & Ori K. "Delay Variation Metrics for Speech Quality Estimation of VoIP". Institute of Electronics, Information and Communication Engineers (IEIC) Technical, 2001.

Mohd Nazri Ismail , "Implementation of Secure Real Time Transport Protocol on VoIP over Wired In-Campus Network Environment" May 2010.

Mohd Nazri Ismail , "Implementation of Secure Real Time Transport Protocol on VoIP over Wireless LAN In-Campus Network Environment" May 2010.

Sharma, A.K "A study of WLAN Roaming and its implications for Voice Services". Seminar on Internetworking Helsinki University of Technology, 2006.

Narbutt, M. and M. Davis, "Gauging VoIP call quality from 802.11 WLAN Resource Usage," In Proceedings of IEEE Symposium on Wireless, Mobile and Multimedia, 2006.

Mohd Nazri Ismail, "Analysis of Secure Real Time Transport Protocol on VoIP over Wired LAN in Campus Environment", (IJCSE) International Journal on Computer Science and Engineering Vol. 02, No. 03, 2010, 898-902.

Andre L. Alexander, Alexander L. Wijesinha, and Ramesh Karne, 2009, "An Evaluation of Secure Real-time Transport Protocol (SRTP) Performance for VoIP", 2009 Third International Conference on Network and System Security.

Sureshkumar V. Subramanian and Rudra Dutta, "Comparative Study of Secure vs Non-Secure Transport Protocols on the SIP Proxy Server Performance: An Experimental Approach", 2010 International Conference on Advances in Recent Technologies in Communication and Computing.

Adomkus, T. and E.Kalvatis, "Investigation of VoIP Quality of Service using SRTP Protocol", Telekomunikacijų katedra, Kauno Technologijos Universitetas, 2008.

André Leroux, Bruno Giguère, M "SONET/SDH Vs Ethernet: Migration and Testing Issues", 2006.