The Analysis of New Video Conference System Based Secure Authentication

Yong-Deug Jung*, Gil-Choon Kim**, and Moon-Seog Jun*
*Dept of Computer, Soongsil University, Seoul City, 156-743, Korea
Tel : 82-2-820-0680, Fax : 82-2-825-7178 E-mail : jungyd@kotra.or.kr , mjun@computing.ssu.ac.kr
**Dept of e-Business, Sungkyul University, Anyang-city, 430-742,Kyonggi-do, Korea
Tel : 82-31-467-8107, Fax : 82-2-467-8067 E-mail : kimid@sungkyul.edu

Abstract: The paper describes the implementation of the video conferencing system using public key infrastructure which is used for user authentication and media stream encryption. Using public key infrastructure, we are able to reinforce the authentication for conference participant and block several malicious hacking while protecting conference control information. The paper shows the implementation of the transportation layer secure protocol in conformity with Korea public key authentication algorithm standard and symmetric key encryption algorithm (RC2, SEED, DES and 3DES) for media stream encryption. The feature of the paper is transportation layer secure protocol that is implemented for protection of information on a user authentication and video conference and the media streaming encryption algorithm also can be envisioned with another block encryption algorithm. The key for media streaming encryption may be safely distributed by the transportation layer secure protocol.

1. INTRODUCTION

The standard for a video conferencing has mainly been developed by ITU-T(International Telecommunication Union Telecommunication Standardization Sector) and/or IETF(Internet Engineering Task Force). Between them, H.323[1], the standard for a video conferencing suggested by ITU-T, is working protocol at the Internet environment which is not guaranteed the QoS. It is also supposed to be adapted to the environment of the next generation mobile communication.

The most important security weaknesses at the H.323 video conference are as follows; First, the authentication for users is required. It is more sensitive at the commercial part in which the qualified users are participating. Second, secrecy of the information on the video conference crucial. You could be attacked in the middle of a video conference. Third, the integrity for the information exchanged between user-to-user and/or user-to-control server must be guaranteed. Without it, intentional communication distortion and service denial attack is possible.

To find solutions to the above mentioned security problems, the paper would address the TLS[12,13] method. There are reservation method & authentication exchange method to provide authentication, confidentiality, integrity and the paper would use the block encryption algorithm, such as RC2[5], DES[6], Triple-DES[7], to keep secrets on the information on the video conference.

The paper embodies user authentication and encryption /decryption of media about H.323 video conference system. At the H.323 system, the controlled information was protected through the implementation of TLS certification which follows domestic standards. The encryption of media was protected through the implementation of the block encryption algorithm such as RC2, DES, Triple-DES, SEED. The secret key for encryption of the media distributed safely in the process of producing call for sharing by TLS. The characteristics of the paper are as follows;

1. TLS was adapted for user authentication and media encryption
2. Public key based certification, which satisfies the standards, was adapted for enhancing the security
3. Analysis on an affected factor in video conference by researching the character of block encryption algorithm.

We will take a look at the video conference and security technology in Section 2. In Section 3, we will describe video conference model and the security protocol for it. In Section 4, the produced outcome and result for each block algorithm will be suggested. The conclusion is in the Section 5.

2. RELATED WORK

2.1 H.323 Video Conference Standard

H.323 is the standard of multimedia communications supporting transmission back-and-forth of the audio and video data on TCP/IP packet network such as Internet. Its strength lies in that it doesn't change the shape of packet like PSTN and provides the standard for interconnection with other networks such as LAN, GSTN, N-ISDN, B-ISDN. It also supports codec standard for audio and video, interconnection, the network independency, platform and application independency, multi-point service, band management, multicasting, flexibility, conference service with other networks, etc.

H.323 defines components, protocol, procedure for providing real-time point-to-point and multi-point multimedia communication on the packet network. It needs four components; terminal, gateway, gatekeeper, Multipoint Control Unit (MCU). Terminal is real-time two way communication supporting end-to-end device. It supports audio, video, and data service.
H.323 audio codec recommend 5.3~64kbit per second of the scope. Pulse Code Modulation(PCM) of method G.711. G.723.1 supports in this paper which is more than quality G.711 that it is supports audio codec scheme which is typical telecommunication network environment. H.261 is through the connection two way method of the loss and non-loss compression method for the more efficient moving picture compression and then 352 x 288 pixel transmit below 30 frame per second as the 64kbps speed. H.261 supports the Common Intermediate Format(CIF) with display configuration of 352,288 pixel. H.263 supports sub-QCIF(12,896), QCIF(176,144), CIF(352,244), 4CIF(702,576), 16CIF(14,081,152). With QCIF formatting, H.263 is reciprocal to H.261.

2.2 Video conference cryptograph

In this section, we will describes the crypto technology for media data protection. Cipher technology where a single person should encrypt texts and also retrieve plaintext from the encryption.
Cipher technology is mainly divided into two parts; Secret key cipher algorithm and Public key cipher algorithm. We call it "Secret key crypto algorithm" when the encryption key and decryption key is the same, and otherwise, "Public key crypto algorithm".

2.2.1 Public Key Algorithm

Public-Key Crypto Algorithm is an algorithm with two different keys for encryption or decryption. In other words, encrypted with one key could be deciphered only with another key. With this feature, Public-Key Crypto Algorithm is also called Asymmetric Crypto Algorithm. Two keys are a relative pair based on a mathematical function. One key is opened for anyone to use it, while the other is kept from the public. Opened one is called the public key and the other is refered to the private key. The process necessary for encrypted communication between sender and recipient using the public key. Sender transmits the ciphertext message to recipient with K_{ub}, a public key of sender B.

$$C = EK_{Ub}(M) \quad (1)$$

Then, recipient B get the original message through decoding his K_{pb} the ciphertext.

$$M=DK_{Pb}(C) \quad (2)$$

Public key Crypto Algorithm provides digital signature function that Secret key Crypto Algorithm can not. Digital signature is a technology where sender A send the ciphered (signatured) text with its personal key. Recipient B shall, then, decode the ciphered text with A's public key.
Since it is only the sender A who has the private key, you can verify that it is A who had ciphered the text with A's public key. This process has the same effect as that we get when sending sealed documents back-and-forth.

Public key crypto system plays important roles in the security protocol for security of E-Commerce. The two most widely used public-key algorithms are RSA and Diffie-Hellman. We will briefly describe its principles in this section. One of the first public-key schemes was developed in 1977 by Ron Rivest, Adi Shamir, and Len Adeleman at MIT and first published in 1978[13,14]. Major Public key crypto algorithms are; RSA(Rivest, Shamir, Adleman) asymmetric cipher algorithm, ElGamal, ECC(Elliptic Curve Crypto systems).

2.2.2 Transport Layer Security Protocol

Transport Layer Security(TLS) are used as cipher algorithm. TLS is an IETF standardization initiative whose goal is to produce an Internet standard version of Secure Socket Layer(SSL). The SSL protocol supports both stream and block encryption ciphers, although the message formats differ slightly. TLS is designed to make use of TCP to provide a reliable end-to-end secure service. TLS not a single protocol but rather two layer of protocols.

The TLS Record Protocol provides basic security services to various higher-layer protocols. In particular, the hypertext transport protocol(HTTP), defined in RFC2068 which provides the transfer service for Web client/server interaction, can operate on top of TLS. Three higher layer protocols are defined as part of TLS: the Handshake Protocol, The Change Cipher Spec Protocol, and the Alert Protocol. These TLS specific protocols are used in the management of TLS exchanges and are examined later in this section[13].
Between any pair of parties applications such as HTTP on client and server, there may be multiple secure connections. There are a number of states associated with each session. Once a session is established, there is a current operating state for both read and write. Handshake Protocol is pending read and write states are created.

3. SCHEME AND IMPLEMENTATION OF VIDEO CONFERENCE SYSTEM

3.1 Video Conference Model

In this section, we will explain the design and embodiment of video conference system applying H.323 standard and TLS. Considering that the communications bandwidth range changes any time in the network environment, we set the program to be operated at the Client for participant to feel the minimum network load. And we also get them transmitted dividely as a multimedia graphic packet unit.
The model shown at the paper is based on the TCP/IP protocol, and organized to hold a conference through the Web server and Image Server. The communication among participants will be possible in Peer-to-Peer way that doesn't use the video conference server. There are two exceptions. One is the communication via server for firewall user, the other is the authentication

procedure when the participant initially enter the conference room.

3.1.1 Peer-to-Peer Based New Video Conference

Peer-to-Peer(P2P) based conference model directly sends all the information exchanged among participants once a conference begins, while it gets help of Server, before the conference. Fig.1 shows the scenario under which Peer-to-Peer(P2P) based conference operates.

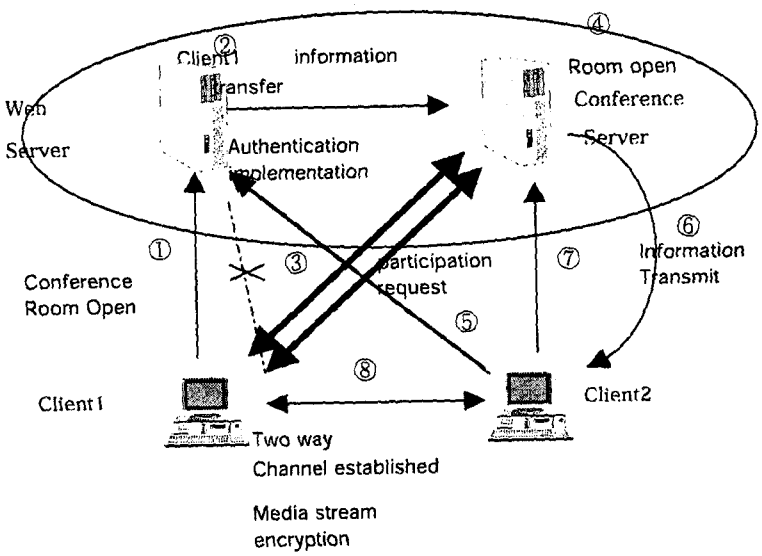

Fig. 1. Peer-to-Peer New Video Conference Model

Step 1 : Client 1 makes request for opening of conference room to the Web Server. At this point, Client sends the previously agreed room number.

Step 2 : Web Server produces the process of Conference Server and sends the information of Client 1 to the process. It also sends the information on Conference server process to the Client 1.

Step 3 : Client 1 ends the connection with the Web server and makes request for the opening of conference room to the conference server. When the conference begins, Client 1 transmits the information capable of communication with other Client, to the conference server. When Client accesses the video conference server for meeting, the TLS adapted authentication is a must.

Step 4 : Conference Server makes the conference room when the room number from web server and client 1 matches.

Step 5 : Client 2 makes request for starting of conference to the web server. At this point, Client sends the previously agreed room number.

Step 6 : Web Server makes sure whether the requested conference room is open and sends the information of Client 2 to the video conference Server. It also sends the information of managing video conference server to the Client 2.

Step 7 : Client 2 inquiries the information of Client 1 to the video conference server process and get response.

Step 8 : Client 2 requests two way channel from the Client 1 and starts the video conference. At this point, the encryption process is adapted about the stream data needed at the video conference.

Peer-to-Peer based model can increase efficiency of network and doesn't need a separate server because it directly sends all the produced data to the participants. However, the organization with the firewall system may need a separate server. We will address it in the following section.

3.1.2 Server Based Video Conference

Server based conference model exchanges the participant's information not only in the conference preparation but during the conference. The procedure of it is shown Fig. 2. Server based conference model has an inefficiency problem, which all produced data during the conference via the Server. However, under the access-deniable circumstances owing to the firewall system, server based conference could be used for the choice of Client.

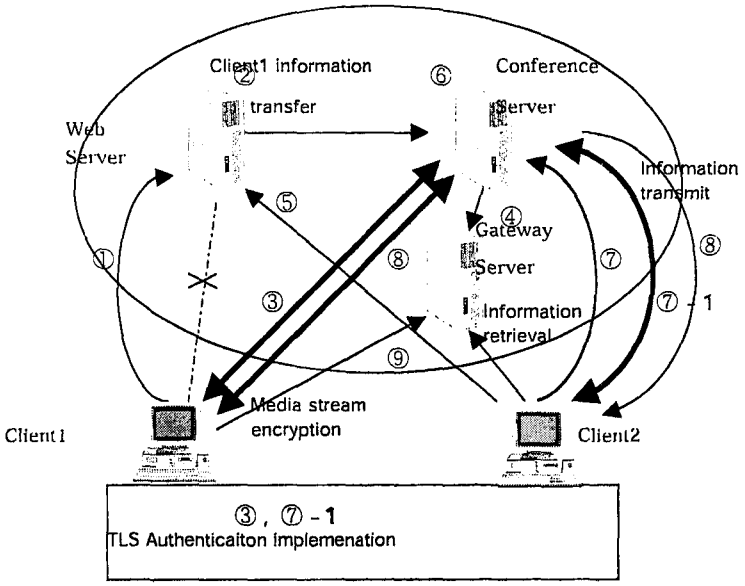

Fig. 2. Server Based New Video Conference Model

Step 1 : Client 1 makes request for opening of conference room to the Web Server. At this point, Client sends the previously agreed room number and the information whether the firewall is used.

Step 2 : Web Server produces the video conference server and sends the information of Client 1 to the server. It also sends the information of video conference server to the Client 1.

Step 3 : Client 1 ends the connection with the Web server and makes request for the opening of conference room to the conference server. When Client accesses the video conference server for meeting, he/she should undertake a TLS-employed authentication process.

Step 4 : Provided that the conference room numbers it gets from web server and the client, Video conference sever opens the room and produces Gateway process that will relay the image information of the participants.

Step 5 : Client 2 requests Web server to start conference, when it relays the pre-agreed room number.

Step 6 : Checking if the room that falls on the requested room number is open, if it is, the Web server sends to the conference serve process the information on Client 2 and to Client 2, the information on video conference process in charge of the room in question.

Step 7 : Client 2 refers the information on Client 1 in video conference server process. 7-1 involves TLS verification.

Step 8 : Video conference server process provides Clients 1 and 2 with information on Gateway Process.

are connected via Gateway, the Gateway process shall relay image information and encode the media data.

Step 9 : Clients 1 and 2 starts conference by requesting connection to the Gateway process. Once the Clients At this time, it sends lists of several information such as the Client-usable algorithm list, public key algorithm list, compression method list. The server, receiving the Client Hello message, decides to adapt it from the list of parameter and send the Server Hello message in response. It also sends the own certification and the electronically signed value of parameter to the Client.

3.2 Applying Security Protocol

Suggested video conference at the paper consists of web server, video conference server, gateway server, and participants. To sum up the security flaws possible among the components, security steps should be taken such as authentication of participants, exchanging control information for the conference, media information among participants during the conference, etc.

So these security defects could be solve by Transport Layer Security Protocol with security certification. It could provides, as addressed in Section 2, a reinforced authentication service as well as that of participants and hardware resources used in conference.

In addition, transport Layer Security Protocol provides the strength that it could be adapted to the other components without distortion. Transport Layer Security Protocol made up of the authentication procedures between Client and Server as shown in Fig. 3. Full Handshake process of Transport Layer Security(TLS) starts when the Client sends Hello message to the Server.

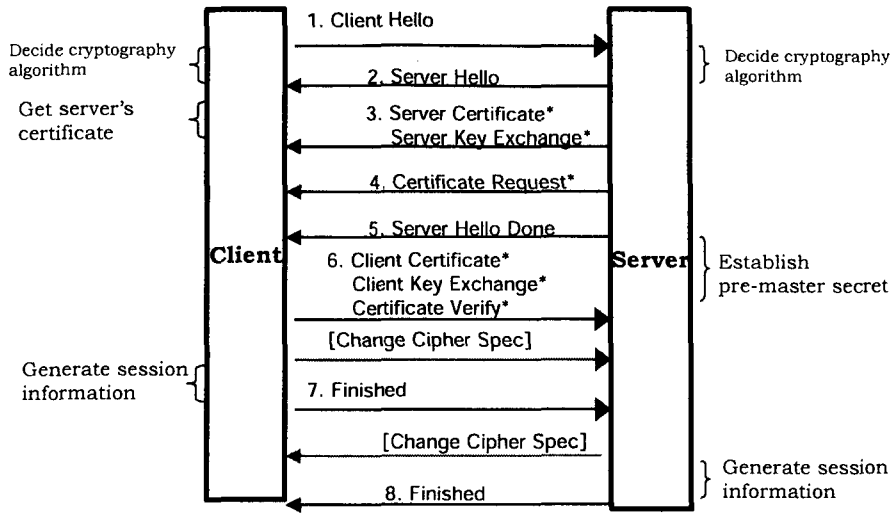

Fig. 3. Transport Layer Protocol Operation

It requests Client certification. Client could certify the Server initially by checking the signature on the electronically signed parameter. After successful completion of signature checking, Client sends the certification to the server upon request. Also, it electronically signs the used message and sends it to the Server. Server checks the message. If it is completed without errors, it is considered a success. Through this process, Client and Server, which produced necessary encryption parameter, are ready to exchange the user data after completing the Handshake procedure.

4. VIDEO CONFERENCE SYSTEM ENCRYPTION EVALUATION ANALYSIS

4.1 Video Conference System Encryption Implementation

In the paper, we used Authentication for Client(AC) module and Authentication for Server(AS) module for authentication of video conference and encryption as shown in Fig. 4. We also applied Cipher message Buffer (CB) for encryption and decryption of media data used in conference.

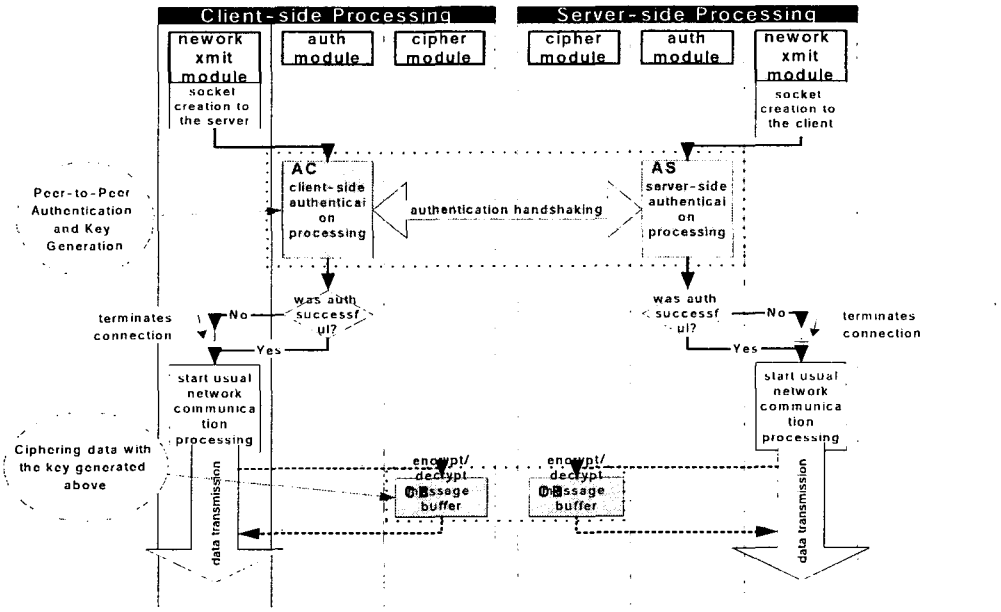

Fig. 4. New Security Module Diagram

Codec for image compression and voice compression using G723.1 and GSM6.10 for PCM way. Average network bandwidth was set as 1.4kb for voice, 4kb~2kb for image, 1kb for other in terms of bandwidth. In case of variation for bandwidth, we use the changed protocol. Image bandwidth is generally transmitted with 32bit. We set mean of bandwidth as a calculation of bandwidth and the values drawn from the mean was calculated as a minimum value. Message Buffer was used as a method of encryption for data and

we could process the media delaying speed as fast as we can. That's why the encrypted video conference system generally slow even though the stream encryption is adapted. For the media simulation for encryption and decryption, as shown on Fig. 6, the age of first 256byte is transmitted from the Client end and Fig. 7 indicates that the decryption with receiving the encrypted data from the Server end.

4.1.1 Media Stream Encryption /Decryption Implementation results

Encryption of media stream consists of common key production by call setup motion, distribution of secret key, the encryption of audio and video stream. The transmitted message stability is related with the user-key relation R. Namely, when the User, Key fair belongs to user the user-key relation R, the message is stable.

$$R = (User, Key) \quad (3)$$

Message encryption by itself can provide a measure of authentication. A message M transmitted from source A to destination B is encrypted using a secret key K shared by A and B. If no other party knows the key, then confidentiality is provided: No other party can recover the plaintext of the message. We assumes two user communicating parties, user A and B, share a common secret key K. It calculates the MAC(message authentication code) as a function of the message and the key : MAC = C_k(M), where M = input message,
C = MAC function, K = shared secret key,
MAC = message authentication code.
To provide authentication, A uses its private key to encrypt the message, and B uses A's public key to decrypt as shown Fig 5. This provides authentication using the same type of reasoning as in the symmetric encryption case : The message must have come from A because A is the only party that possesses KR_a and therefore the only party with the information necessary to construct ciphertext that can be decrypted with KU_a. Again, the same reasoning as before applies: There must be some internal structure to the plaintext so that the receiver can distinguish between well-formed plaintext and random bits.

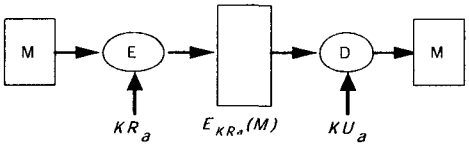

Fig. 5. Public-key encryption:authentication and signature

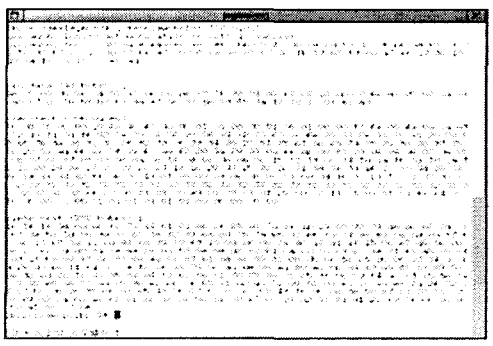

Fig. 6. media stream encryption result

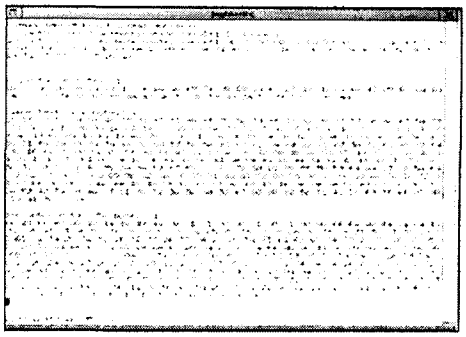

Fig. 7. media stream decryption result

4.1.2 Audio Stream Encryption Results

Encrypted audio and video data of video conference data can not be seen or heard without a valid key. Especially, the video screen is seen with full of noises and more important audio data can also be heard only the each terminal. Fig. 8,9 shows the case when the wrong key is used.

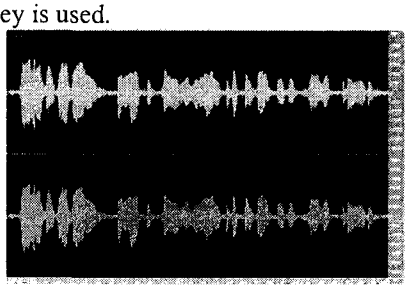

Fig. 8. not-decrypted voice modulation

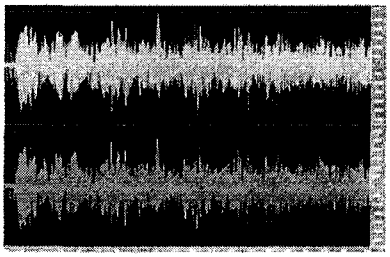

Fig. 9. using normal key voice modulation

Fig. 8 indicates the transformation from sound wave transformer with adequate key. We can hear the precise sound and contents from it. Fig. 9 indicates the transformation from sound wave transformer without adequate key. We can't hear the precise sound and contents from it and the wave is distorted. Fig. 10, 11 is the Histogram of comparison on audio wave. It shows that normal audio wave and abnormal one have a great difference from the shape of wave.

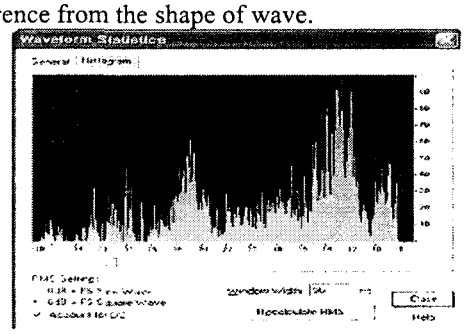

Fig. 10. Conference participant of normally voice modulation

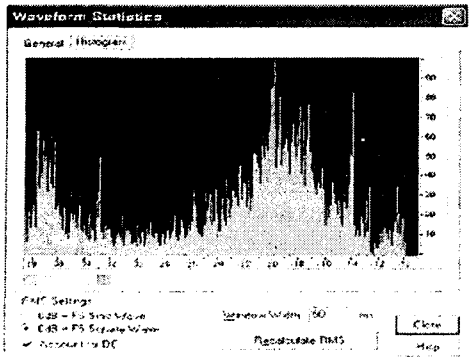

Fig. 11. conference participant of abnormally voice modulation

Table 1. difference between normal decryption and abnormal decryption

	Normal decryption	Abnormal decryption
Min Sample Value:	-20043	-25867
Max Sample Value:	17351	25275
Peak Amplitude:	-4.27 dB	-2.06 dB
Possibly Clipped:	0	0
DC Offset:	0	0
Minimum RMS Power:	-38.88 dB	-37.1 dB
Maximum RMS Power:	-10.1 dB	-8.35 dB
Average RMS Power:	-20.58 dB	-16.05 dB
Total RMS Power:	-18.87 dB	-14.64 dB

Table 1. shows the difference between two data. We can confirm that the all data, for instance the maximum noise, is not the same so we can't say that the file is the same. Therefore, although the unauthorized users access to the video conference system through sniffing or hacking, they can't interfere with the perfectness of data communication.

4.2 Video Conference Stream Encryption Results

Fig. 12 illustrates how to use video conference. Data between users are communicated as encrypted, and the vision is very clear. In Fig. 13, however, captured video conference by sniffing encrypted room number and ID. As a result, we can not hear or see anything because of encrypted message data. Especially in case of picture, the shape is full of noise and existing users can check who is accessing illegally, watching who the intruders are. If the intruder encrypted the data, the existing user can not verify who they are, but this is impossible as they do not share encryption key.

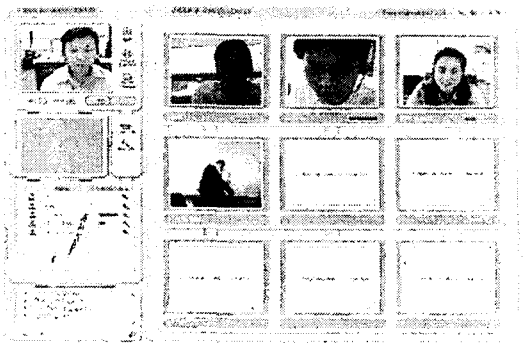

Fig.12. normally video conference

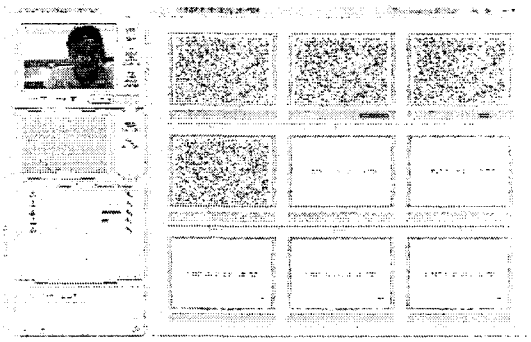

Fig. 13. abnormally video conference

4.3 Evaluation

Table 2 shows the result of measurement of time delay in media transmission both when the video conference data exceeding 5MByte in size were not encrypted and when the data were encrypted with individual symmetric keys. The test was conducted with a computer with four CPU 2Ghz and 512M flash memory as an operating server. Simultaneous video conference was taken place with users who have basic configuration PC from 16 places where network speed exceeds 128 KB/sec.

As Table 2 shows, we found that it took less time to encrypt when we use the RC5 and DES algorithm than other cipher algorithm. However, as Fig. 14 shows, it took only small amount of time to encrypt, and because there is little difference between encrypted time and not-encrypted time, the users can hardly distinguish the difference between the two processes. Supposed Peer-to-Peer based video conference system should not affect the main server according to the experiment. Assuming that firewall is set up at the user end, assessment was done upon the server function when server-based video conference system is in operation and when all 16 users are on the video conference system.

Also, kotra-control daemon, which is a video conference server daemon processing, memory use grew by about 0.1% and CPU use increased by 0.3%. Therefore, the video conference server proposed in this paper can handle server access by more than 1,500 users at a time, utilizing the least of computer resources during the encryption/decryption process.

Table 2. 5Mbyte media transmitting time(unit : micro second)

Algorithm	Total time	transmit time	encryption time	increasing rate(%)
Not encrypted	42,452,264	42,452,264	0	0
3DES_3KEY	45,207,904.70	42,450,358.70	275,754.60	0.649563943
3DES	45,222,272.40	42,452,704.40	276,956.80	0.65239583
DES	43,554,251.30	42,450,883.30	110,336.80	0.259907929
IDEA	43,839,485.50	42,454,157.50	138,532.80	0.326326059
RC2	44,951,342.40	42,460,974.40	249,036.80	0.586627842
RC5	43,244,046.50	42,455,795.40	78,825.11	0.185679402
SEED	43,484,614.70	42,451,300.70	103,331.40	0.2434061

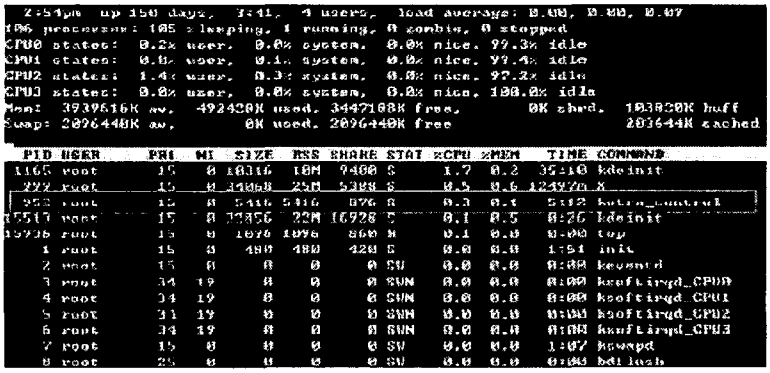

Fig.14. video conference server capacity display

5. CONCLUSION

In this paper, we applied Transport Layer Protocol for encryption algorithm of public authentication standard security protocol. For transport security protocol, we applied public key-based certificate that satisfies Korea public authentication standard. Video conference system controlling information can be protected by applying transport security protocol. This also secured secrecy and perfectness of user authentication and controlling information.

We have applied RC2, DES, Triple-DES, SEED block cipher algorithm for the encrypted implementation of media stream and analyzed each of encryption algorithms' effect on video conference system. As a result, it strengthened confidentiality during the encryption/decryption processing compared with which is insignificantly difference before and after in processing speed and transmitting latency time of the media stream packet in the shared memory using. Therefore, in spite of added function, it was not affected by the processing speed difference in the authentication and encryption/decryption. In addition, we have solved a problem of the secret key distribution that is during of processing the media stream encrypt using a transport layer security protocol.

Reference

[1] H.Schulzrinne, "Simple Conference Invitation Protocol", Internet Draft, IETF, Feb 1996.

[2] Steven McCanne, Van Jacobson, "vic : a flexible framework for packet video", ACM Multimedia, November 1995.

[3] E.Rescorla, Diffie-Hellman Key Agreement Method, IETF RFC 2631, 1999.

[4] Federal Information Processing Standards Publication. Announcing the Advanced Encryption Standard(AES), 2001.

[5] ITU-T Recommendation H.323, Visual Telephone Systems and Equipment for Local Area Networks which Provide a Non-Guaranteed Quality Service(ver4), 2000.

[6] Jacobson, V. and McCanne, S., "Visual Audio Tool", Lawrence Berkery Laboratory.

[7] ITU-T Recommendation H.225.0, Media Stream Packetization and Synchronization Non-Guaranteed Quality of Service LANs(ver4), 2000.

[8] L.Berc, W. Fenner, R.Frederick, S. McCanne, "RTP Payload Format for JPEG-compressed Video", RFC 2035, October 1996.

[9] ITU-T Recommendation H.245, Control Protocol for Multimedia Communication(ver8), 2001.

[10] ITU-T Recommendation H.235 Security Encryption for H-series Multimedia terminals(ver3), 2001.

[11] K.Kelly and J.Mark, "Distributed multipoint conferences using SIP" Internet Draft, Internet Engineering Task Force, Mar.2000

[12] R.Rivest, A Description of the RC2(r) Encryption Algorithm, IETF RFC 2268, 1998.

[13] William stallings, Cryptography and Network security. Pretice Hall, 1998.

[14] S.A.Thomas, SSL&TLS Essentials : securing the web Wiley, 2000.

[15] RFC 2246 The TLS Protocol Version 1.0. T. Dierks, C Allen. January 1999.

[16] RFC 3546 Transport Layer Security (TLS) Extensions S. Blake-Wilson, M. Nystrom, D. Hopwood, J. Mikkelsen, T. Wright. June 2003.

[17] A. and B.Author, "Paper Title" ICASE Journal, vol. 1. no. 1., pp. 123-126, 1996.