INTERNATIONAL JOURNAL OF NETWORK MANAGEMENT
Int. J. Network Mgmt 2001; **11**:147–159 (DOI: 10.1002/nem.398)

Security-enhanced packet video with dynamic multicast throughput adjustment

By Han-Chieh Chao*, T. Y. Wu and Jiann-Liang Chen

In recent years, the Internet population has increased at an explosive rate. Many problems exist because Internet packets are not encrypted and the bandwidth is not large enough. In this research, we propose a datagram encryption technique and dynamic bandwidth throughput adjustment. Security is enhanced using secret keys selected from a Key-Database. Not only is a double encryption tunnel offered but also the whole plaintext can be encrypted more than one key depending on the encryption block size chosen.

Introduction

Video multicast distribution is an important component of many existing and future networked services. Today's Internet lacks adequate support for quality of service (QoS) assurance, which makes the transmission of real-time traffic (such as video) challenging.[1] Many of these problems exist because Internet packets are not encrypted and bandwidth is not large enough. An Internet without a security infrastructure in place is vulnerable to several types of attack. Internet use continues to increase dramatically along with the variety of data exchanged over computer networks.[2,3]

Videoconferencing is becoming part of distributed systems. Many distributed applications require information exchanges over insecure public channels. Private exchanges require protection from eavesdroppers. Secure information exchanges are a necessity in distributed systems. Encryption and decryption provide the basic technology for building secure systems. There are two encryption methods: secret key encryption and public key encryption. The decryption available from currently well-known public key schemes is slower than that in secret key schemes. Under security systems, a single key is used for both encryption and decryption and only authorized users possess this key.[4] We present a security-enhanced secret key encryption method using a Key-Database and traditional encryption. This system encrypts plaintext using random keys selected from a Key-Database. The brute-force approach can then be avoided for it is too time consuming.

Who handles retransmission and how retransmissions are processed are key distinguishing factors among reliable multicast transport protocols.

At the same time, we try to avoid overburdening the sender with control traffic and retransmission duties. Who handles retransmission, and how retransmissions are processed are key distinguishing factors among reliable multicast transport protocols. The source is ultimately responsible for retransmissions, but that doesn't necessarily mean that the source must be directly involved in each

Han-Chieh Chao, T. Y. Wu and Jiann-Liang Chen teach at the National Dong Hwa University, Hualien, Taiwan, ROC.

Correspondence to: Dr. Han-Chieh Chao, Department of Electrical Engineering, National Dong Hwa University, Hualien, Taiwan, ROC.

Contract/grant sponsor: National Science Council of Taiwan; Contract/grant number: NSC 89-2119-E-259-002.

retransmission. Most of the previous approaches support real-time video transmissions in integrated services networks that rely on traditional preventive congestion control. Feedback control mechanisms are already used in the Internet to control non-real-time traffic sources. Feedback mechanisms for video sources have also been proposed for networks with variable capacity channels such as the Internet. Using multicast transmissions can reduce the datagram flow within WAN.[5–7] Many early congestion control methods involved adjusting the video quality and data rates over a relatively wider range. Real-time video and IP/TV adjust the bit rate to change the frame encoding to comply with the bandwidth. Traditional methods generate low-quality images that are sent to all subscribers regardless of whether their network nodes have reception congestion or not. The proposed method will only send low-quality encoded images to the nodes that are actually congested. The host server load is thus reduced since it is not necessary to regenerate the entire picture. The picture frames will drop out according to the priority assigned by the host at the beginning. The Key-Database encryption method and frame priority estimation are processed at the same time. The multicast router then controls the bandwidth accordingly.

The paper is organized as follows. The proposed security enhancement with Key-Database is described in the next section. This system contains the International Data Encryption Algorithm (IDEA) and Data Encryption Standard (DES) conventional encryption methods. The Key-Database enhances security using IDEA and improves the tunnel topology. The proposed dynamic throughput adjustment scheme is introduced in the third section. This scheme begins by considering multicast videoconference throughput adjustment over the Internet. The major issue is multicasting over the Internet. The question then arises about videoconference standard H.263. The main objective is dynamic throughput adjustment. In the following section, the experimental results are listed. Conclusion is presented in the final section.

Security Enhanced with Key-Database

Any cryptographic primitive, such as a block cipher or a digital signature algorithm, can be thought of in two very different ways. IDEA is a product block cipher that resists all current publicly known forms of crypto-analysis. We describe a technique for enhancing the IDEA symmetric cipher using Key-Database. There are several conventional encryption algorithms available over secure network systems. IDEA cryptography was selected from several existing security algorithms. Table 1 compares several cryptographic systems that are available currently. IDEA has several important advantages. Because encryption algorithms can be used over an open source, they are better than other conventional encryption algorithms.

Compared to public-key encryption schemes, RSA, the structure of encryption algorithms is very complex. It is very difficult to explain RSA or similar algorithms in details. We propose using conventional encryption over videoconference, because usually more than one user will join the videoconference. If a datagram must be sent to newly joined users, public-key encryption is used for every user of the datagram. When there are

	DES	IDEA	RC5	Skipjack
Designer	NSA	Lai, Massev	Ron Rivest	NSA
Development	1977	1992	1994	1993
Data (bits)	64	64	64	64
Encrypted key	56	128	Variable	80
Round number	16	8	Variable	32
Algorithm (open source)	No	Yes	Yes	No

Table 1. Comparison of several conventional encryption algorithms

many participants, using a nontraditional scheme, the datagram must be encrypted as many times as the total participant number. Conventional encryption only needs to encrypt the datagram once. This definitely decreases the source-host load. The time required to encrypt a video file is approximately one quarter of the total video display time. That is, for a one-hour long, 15 minutes are needed to encrypt the entire file.

—Key-Database—

A Key-Database scheme is presented to enhance IDEA encryption algorithms without altering the original algorithms. The original IDEA encryption algorithm encrypts datagrams using one key. When the ciphertext flows from the source to a destination, it is vulnerable to an interception. An unauthorized party can capture the encrypted data in a network and use any method, such as the brute-force approach, to decipher the encrypted text. On average, half of all possible keys must be tried to achieve a brute force decryption. Table 2 shows how much time is involved for the various key sizes.

Table 2 shows the results of a 56-bit key size with 10.01 hours for decryption. The required cost and time is substantial. Although such a scheme, with a long key, presents formidable crypto-analysis difficulties, it can be broken with sufficient ciphertext using known or probable plaintext sequences, or both. The **one-time pad** is unbreakable. It produces a random output that bears no statistical relationship to the plaintext. Because the ciphertext contains no information about the plaintext, there is simply no way to break the code.

A Key-Database is proposed to replace the single-key scheme.[8] This approach involves using random keys from the Key-Database, producing a 65535 random key selection. The encryption process blocks are shown in Figure 1. Using our approach, the videoconference leader decides how often to replace a key from Key-Database. This conforms to the 'One–Time Pad'.

Figure 2 shows Key-Database IDEA encryption processes. We may recall that 4064 bits of plaintext is encrypted using the block size of 500 bits (approximately using $4064/500 = 8$ different keys), it only takes 0.038 second to complete the whole video encoding and encryption. The proposed method has all of the advantages of IDEA. The time required for changing keys is only 0.935 μs obtained through experiments. This places little overhead on the whole process compared to 38 ms.

—Building a Double Encryption Tunnel—

This approach can also be applied to firewalls and virtual private networks (VPNs).[9] The IPSec supports these features and is mandatory for IPv6 and optional for Ipv4. In both cases, the security features are implemented as extension headers that follow the main IP header. The extension header for authentication is known as the authentication header (AH). The extension for encryption is known as the Encapsulating Security Payload (ESP) header. We would like to focus attention on the tunnel mode. The tunnel mode provides protection for the entire IP Packet. To achieve this, after the AH or ESP fields are added to the IP packet, the entire packet, plus the security field, is treated as the payload for the new 'outer' IP packet with a new outer IP header. The entire original, or inner, packet travels through a 'tunnel' from one point in an IP network to another. No routers along the way are able to examine the

Key size (bits)	Number of alternative keys	Time required at 1 encryption/μs	Time required at 10 M encryption/μs
32	$2^{32} = 4.3 \times 10^9$	$2^{31}\mu$ s = 35.8 s	2.15 ms
56	$2^{56} = 7.2 \times 10^{16}$	$2^{55}\mu$ s = 1142 years	10.01 hours
128	$2^{128} = 3.4 \times 10^{38}$	$2^{127}\mu$ s = 5.4×10^{24} years	5.4×10^{18} years

Table 2. Average time required for an exhaustive key search

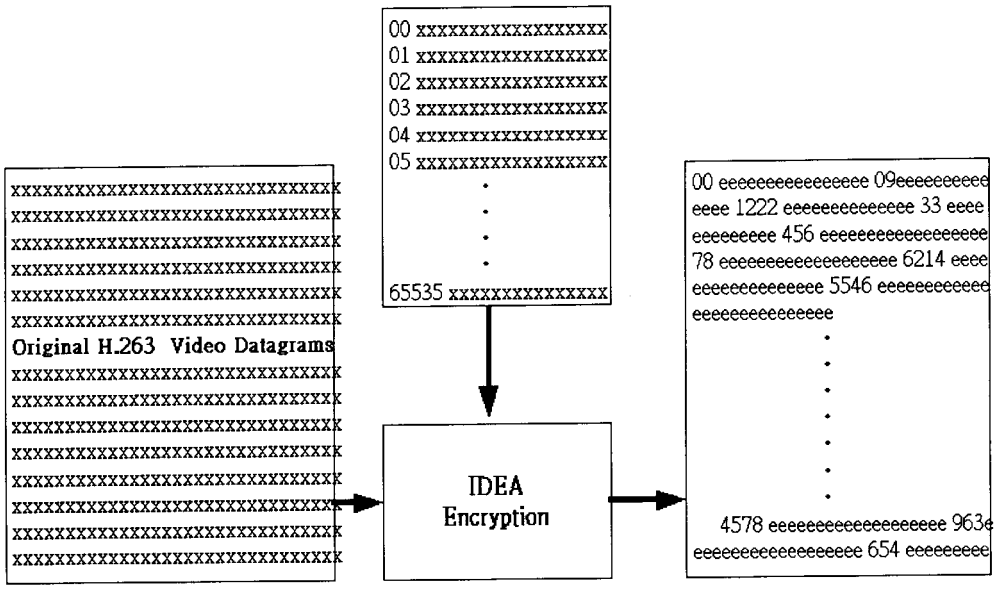

Figure 1. Key-Database encryption block

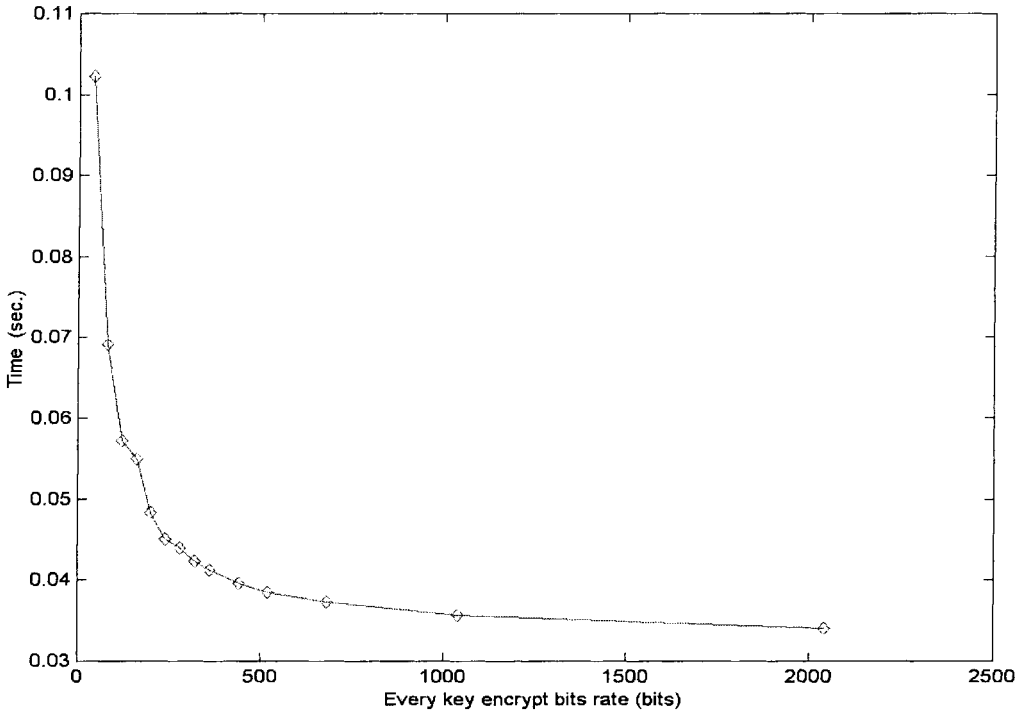

Figure 2. Corresponding time for encryption vs data size of the keys used

inner IP header. Because the original packet is encapsulated, the new, larger packet may have totally different source and destination addresses, adding to the security (Figure 3).

Here is an example of how the tunnel mode IPSec operates. Host A on a network generates an IP packet with the destination address of host B on another network. This packet is routed from the

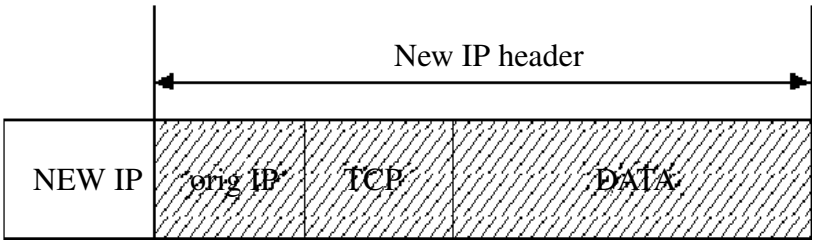

Figure 3. Tunnel mode

originating host to a firewall or secure router at the boundary of A's network. The firewall filters all outgoing packets to determine the need for IPSec processing. If this packet from A to B requires IPSec, the firewall performs IPSec processing and encapsulates the packet in an outer IP header. The source IP address for this outer IP packet is this firewall, and the destination address may be a firewall that forms the boundary to B's local network. This packet is now routed to B's firewall, with intermediate routers examining only the outer IP header. At B's firewall, the outer IP header is stripped off, and the inner packet is delivered to B.

In the encryption method over a traditional firewall and VPN, the ciphertext will be translated into plaintext when it goes through a router. The datagram will be in plaintext when it is transmitted from the router to the destination host. Any participant in the Intranet can access that datagram without authorization.

We noted earlier that the Key-Database acts as the host encryption process and performs double encryption. Figure 4 shows the double encryption. The host initially encrypts the data before transmission. The router then encrypts the entire packet under the original IP header. This double encryption enhances the security from the source end to the destination.

The Internet provides a challenging environment for transporting real-time compressed digital video.

Dynamic Throughput Adjustment

The Internet provides a challenging environment for transporting real-time compressed digital video. Real-time video is generated at the source in a periodic fashion but at a variable bit rate.[10,11] I frame periodicity must be preserved to consider as the lost playout. To accommodate these playout requirements, the network delay jitter must be small. Buffering at the receiver can help absorb some delay jitter up to a limit imposed by the maximum buffer availability. In the case of interactive applications, the end-to-end delay can be reduced.

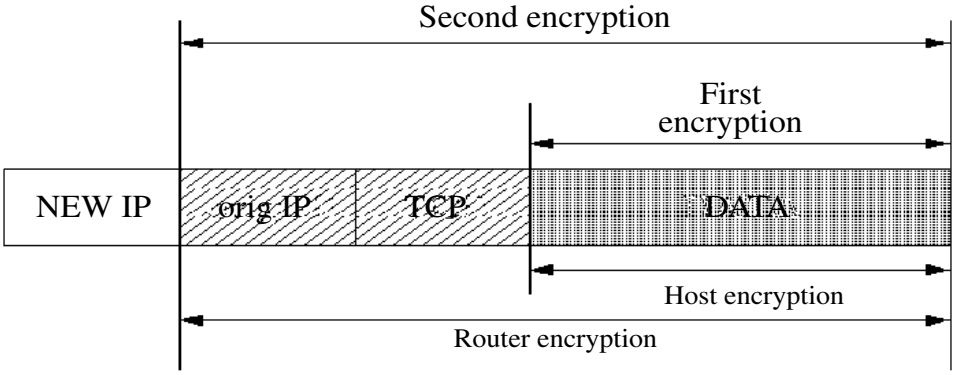

Figure 4. Tunneling mode double encryption

Real-time video has a limited tolerance for random loss within the compressed digital video stream. Excessive losses resulting from network congestion can cause significant degradation of the perceived quality of the decoded video at the receiver.[12,13]

Before referring to our research, we must discuss the H.263 low-bit rate encoding method. Recall our earlier example of the H.263 encoding method (Figure 5). There are three types of encoded pictures (frames) in an H.263 stream: I (intra-coded), P (predicted), and B (bi-directionally predicted) frames. We would like to emphasize the frame encoding for both the inter- and intra-frames.[14,15] The intra-frame is very important because it plays a role in producing later frames. This is the very core of the problem. We will begin by considering a frame with different weights.[16–17] The purpose of this research is to propound a method to drop off frames by priority. In multicast video transmission, the bandwidth is efficiently used. When congestion occurs, congestion control is initiated in accordance with packet priority for the Mrouter. When a frame is skipped, the receiver will repeatedly display the previous frame (Figure 6). Using the proposed control algorithm, we can reserve the high-priority frames to ensure quality.

In this study the emphasis is on encryption and congestion control. Figure 7 shows our research

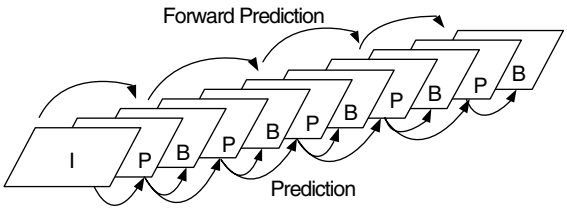

Figure 5. PB frame mode

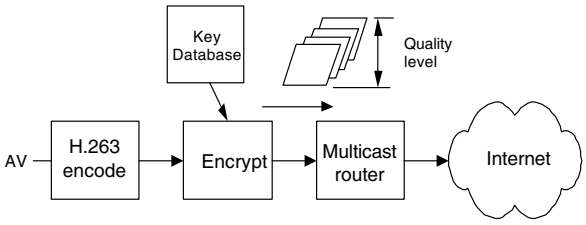

Figure 7. The block diagram of encryption and congestion control

block diagram. Our concern is to consider the datagram for H.263 encode. We proposed an encrypted block diagram by using a Key-Database encryption method and estimated the frame priority at the same time. Priority frames are determined by the number of encoded blocks and the bit rates. The multicast router can then control the bandwidth accordingly.

RTP (Real-time Transport Protocol) is the Internet standard protocol that provides end-to-end network transport functions for real-time data transmission over multicast or unicast network services. It consists of data and control segments. The RTP data segment is an application-layer framing protocol that provides support for applications with real-time properties (e.g. timing reconstruction, loss detection, security, and content identification). The control segment, called Real-time Transport Control Protocol (RTCP), monitors the data delivery in a manner scalable to large multicast networks and provides minimal control and identification functions.

RTP does not guarantee quality of service (QoS) to applications, nor does it provide for resource reservations. Of course, protocols such as the resource reservation setup protocol (RSVP) and Integrated Services over Specific Link Layer

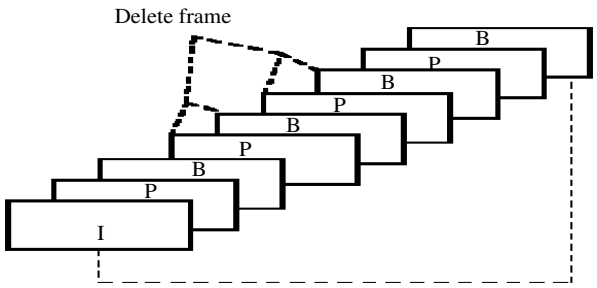

Figure 6. Low-priority frame skipping

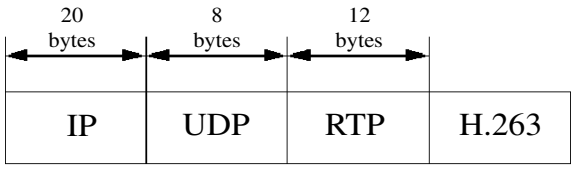

Figure 8. Packet format for RTP

(ISSLL) standards could be used in conjunction with RTP and QoS-based routing to provide certain traffic special handling. These extra features are outside the scope of RTP. Figure 8 shows the packet format for RTP.

The main reason RTCP exists is to provide senders with feedback regarding the quality of their data distribution. This feedback is essential for RTP to succeed in its role as a transport protocol. RTCP's feedback mechanism is analogous to the flow and congestion control functions in other transport protocols. In the case of IP multicast-based sessions, this reception feedback enables an observer or perhaps even a third party to the session, such as an Intranet manager, to monitor a session's status. Any problems may be detected and corrected while the session continues. RTCP's Sender and Receiver Reports (SRs and RRs) constitute its feedback mechanism. RRs and SRs are just 'recommended' for RTP sessions in general, but they are mandatory for multicast sessions.

RTP and RTCP are of particular interest in our discussion since they can be used to provide the quality-of-service feedback from receivers typically required in adaptive video multicast protocols. In the proposed method, RTP and RTCP are used to provide the parameters to preserve the desired QoS.

The priority mechanism for the proposed method can decide the quality of a specific frame. The priority is set, based on the frame intra- or inter-number. The sequences are in QCIF format (i.e. 176 by 144 pixels), and can be set 99 blocks apart, as shown in Figure 9. From the H.263 codec frames, the intra- frame is very important for producing later frames. The PB frames only transmit blocks from the previous frame. The PB frame priority is determined according to the total number of different blocks. The greater the number of blocks, the higher the priority. Therefore the Mrouter can look for the mark within the IP head and determine which frames to skip. We would like to emphasize that both priority setting and the encryption datagram were processed at the same time.

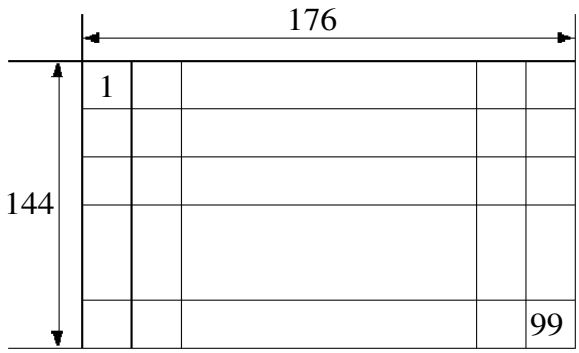

Figure 9. QCIF blocks motion

The resume codec method and proposed method will be compared in this section. The quality of a frame under congested conditions will be considered. When networks have congestion, the source-host will resume codec. All participants within the videoconference will receive low-quality frames, as with real video and IP/TV. When networks have congestion under the proposed method, participants with enough bandwidth will receive the original PSNR sequence. Participants without enough bandwidth will receive low PSNR sequences.

—Experimental Results—

We used multicast topology model to demonstrate the proposed method (Figure 10). The simulation topology includes seven multicast routers and six videoconference participants. In the past the quality of a frame was decided using the objective fidelity measure, PSNR (Peak Signal to Noise Rate).

Those sequences were adopted in this experiment. Three sequences, Claire [Figure 11(a)], Stefan [Figure 11(b)] and foreman [Figure 11(c)] were used for evaluating the proposed methods. Those sequences are in QCIF format (i.e. 176 by 144 pixels) and have different motions. Figures 12(a), (b) and (c) show that when congestion occurs, we can observe that the proposed methods have better congestion control than the no-control method. The no-control method may skip I frame information, which is an important frame for H.263

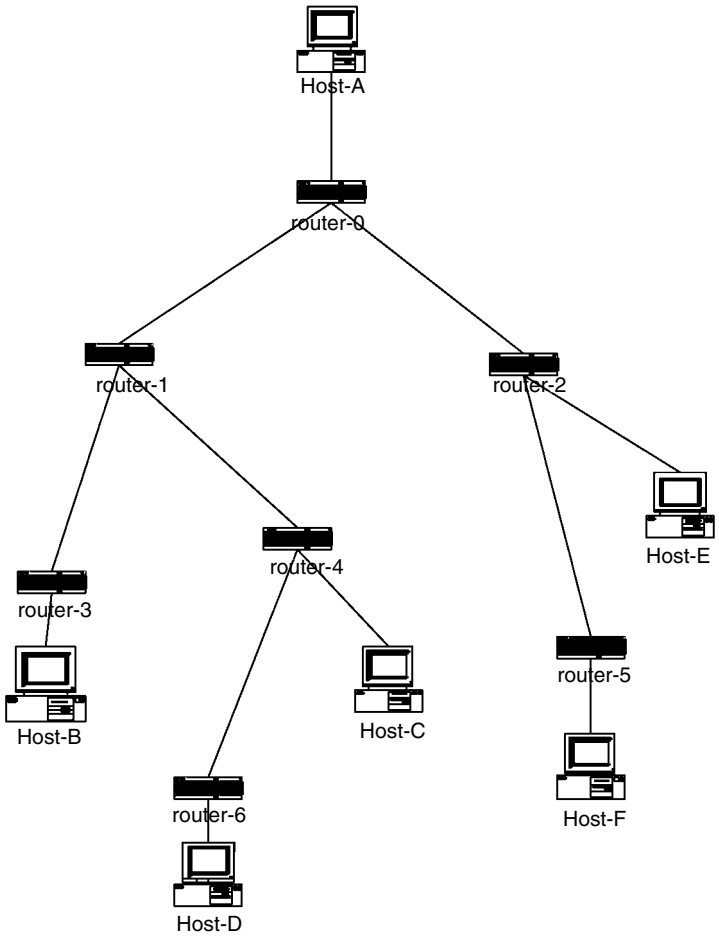

Figure 10. Simulation topology

Figure 11. Sequences adopted in the experiment: Claire (a), Stefan (b) and Foreman (c)

standard. In Figures 12(b) and (c) the PB frames bit rates are similar to I frame for they contain a lot of action within the pictures so that both curves for control and no-control are close together.

Figures 13(a), (b) and (c) show the frame numbers played at the congestion terminal. The diagram illustrates that under the same bandwidth and no-control, the no-control method can skip

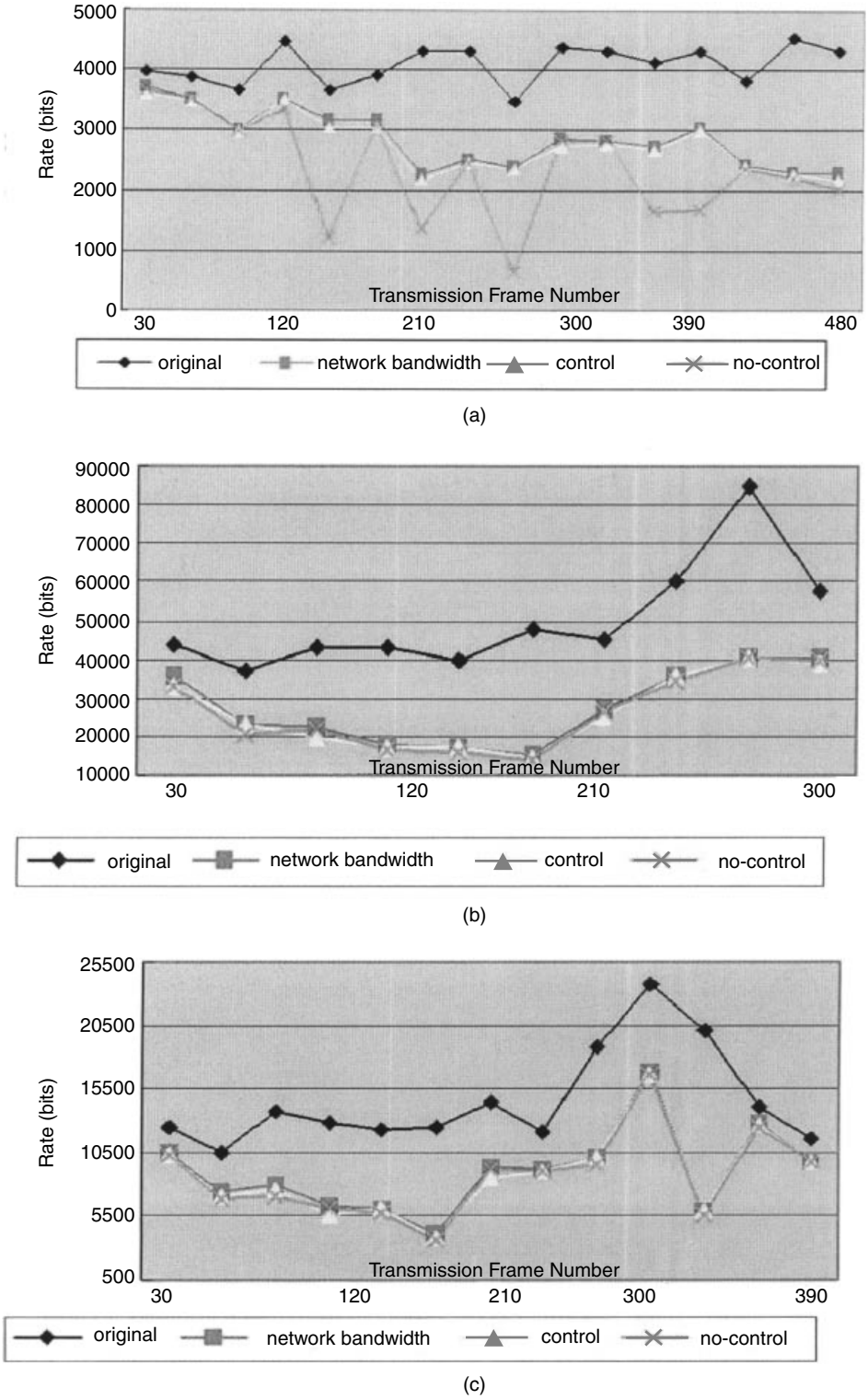

Figure 12. (a) Claire, (b) Stefan, (c) Foreman

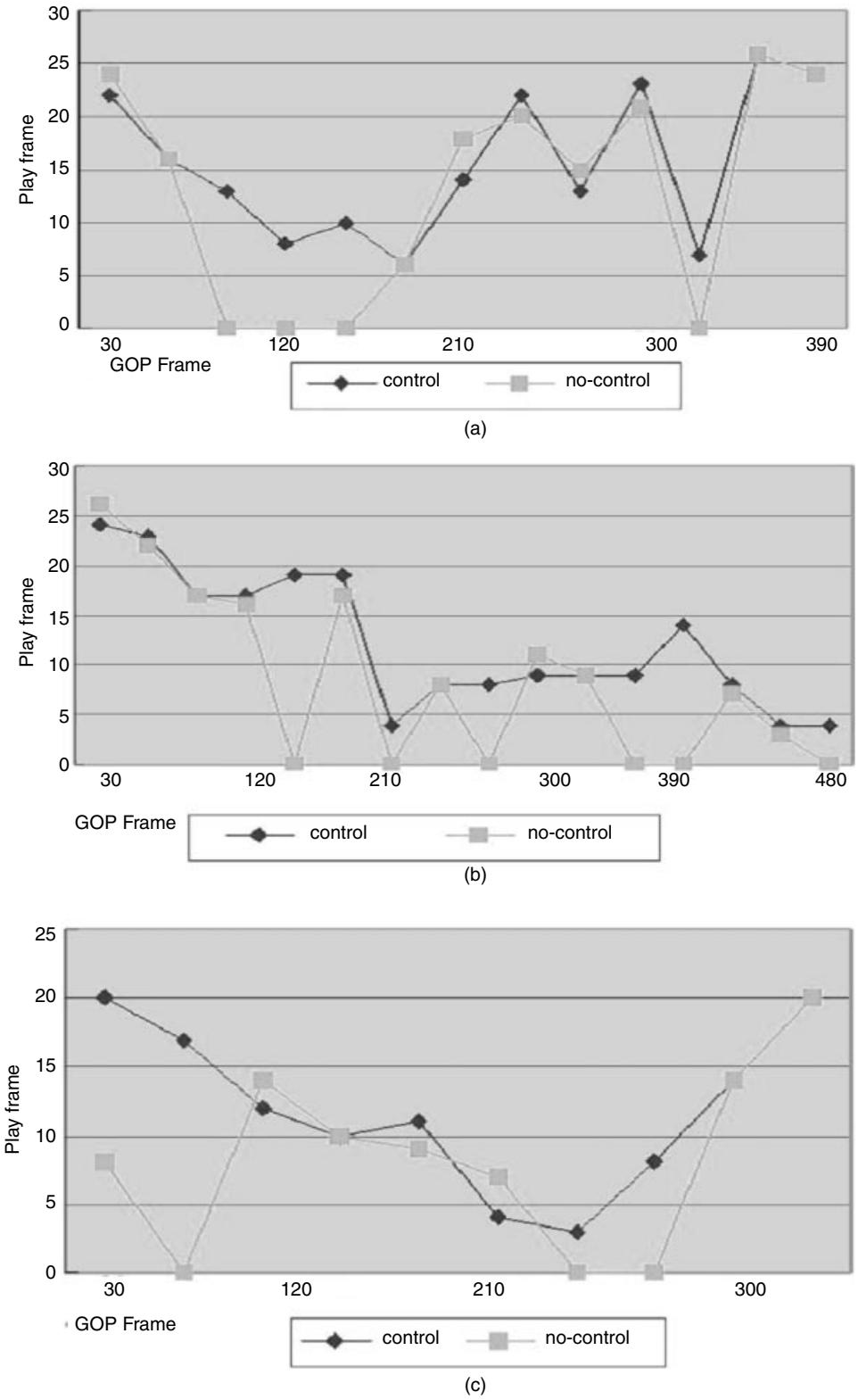

Figure 13. (a) Claire, (b) Stefan, (c) Foreman

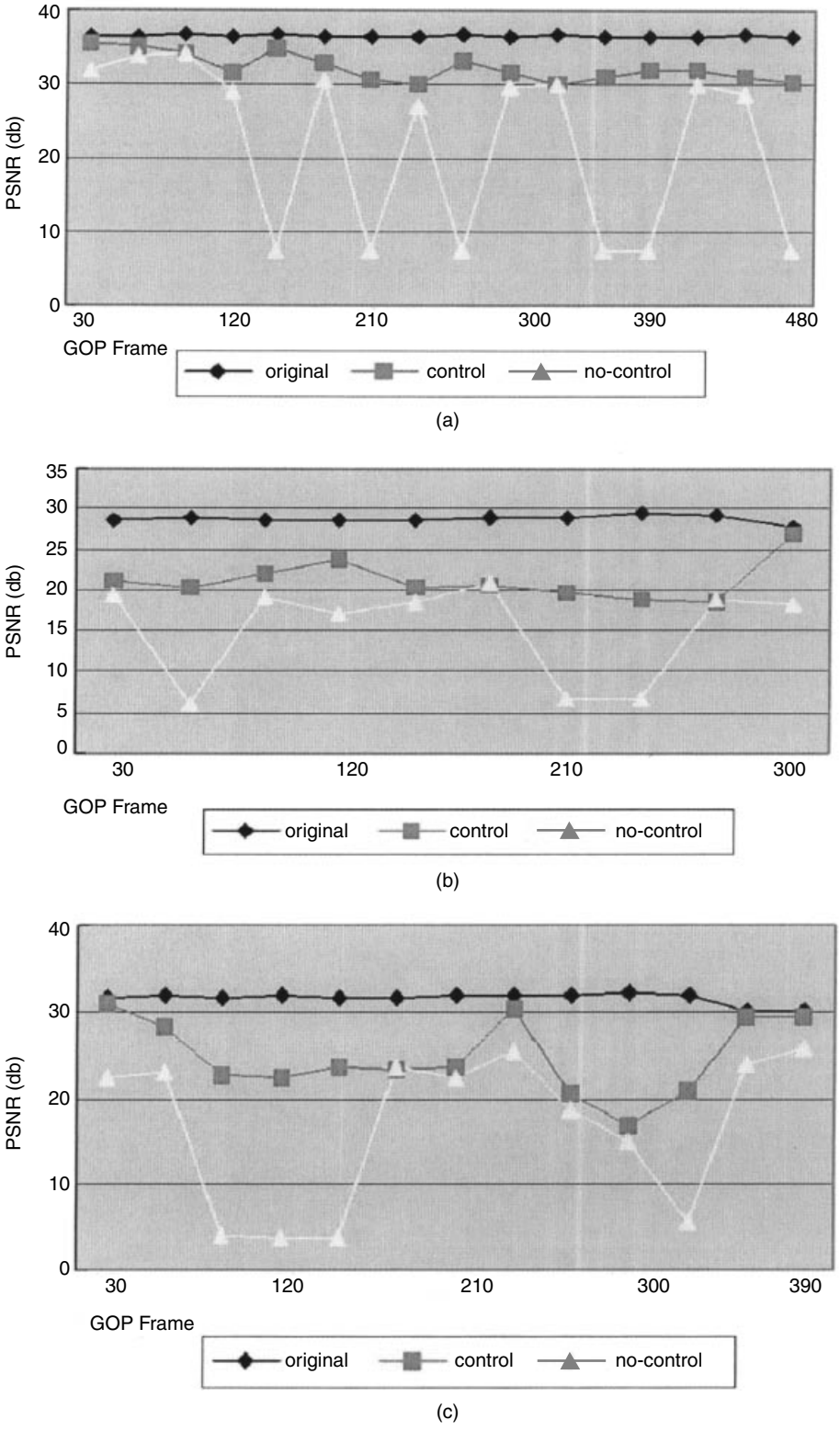

Figure 14. (a) Claire, (b) Stefan, (c) Foreman

the important I frame. The receiver then cannot generate the GOP frame, because the GOP frame is referred by the I frame. This figure indicates that more than one sequence was lost at the no-control terminal. Figures 14(a), (b) and (c) show adjacent frames' PSNR values according to bandwidth. These figures show that the proposed method has gained better PSNR than without control.

The three examples, Claire, Stefan and Foreman, contains different degrees of quality. Table 3 shows the average PSNR for each sequence. The human eyes can barely accept video quality greater than 20 dB. Claire has higher PSNR than Foreman and Stefan with congestion control, because Claire has less action and a simple background. The PB frame contains a lower bit rate than the other two. However, with high action and a complex background, the proposed method can still provide acceptable PSNR values.

	Original	Control	No-control
Claire	36.526 dB	32.261 dB	21.844 dB
Stefan	28.75 dB	21.124 dB	15.144 dB
Foreman	31.593 dB	24.779 dB	16.667 dB

Table 3. A PSNR comparison of Claire, Stefan and Foreman under the proposed scheme

Conclusion

In this paper, we proposed an enhanced packet video security method with dynamic multicast throughput adjustment. The method uses the H.263 system standard, so it can be implemented in existing videoconference systems easily, without needing to change the hardware. Our IDEA enhanced security method encrypts plaintext using a random key form Key-Database and can effectively prevent the brute-force attack.

Many of the early methods regarding congestion control involved adjusting the video quality and data rates over a relatively wider range. We can say with certainty that using multicast to transmit datagrams is more efficient for bandwidth usage. Other methods can only transfer one flow of the same quality to every participant at the same time while ours is not. The main difficulties in video multicast feedback control are feedback implosion and heterogeneity in the network. The proposed Mrouter congestion control can balance the load. It only supplies low-quality sequences to the congested end. Users without congestion can still receive pictures with the original quality. The dispersed control load can be easily handled by Mrouters. This can further reduce the source server load. The response to the transmission rate is also much faster than other methods under congestion. The proposed method uses the same amount of time to complete both the encryption and priority determination for every frame. It is fast, low cost and transparent to both clients and servers.

We can say with certainty that using multicast to transmit datagrams is more efficient for bandwidth usage.

Acknowledgement

This work is partially supported by National Science Council of Taiwan, ROC, under grant number NSC 89-2119-E-259-002.

References

1. Xue Li, Ammar MH, Sanjoy P. Video Multicast over the Internet. *IEEE Network* March/April 1999.
2. Al-Salqan YY. Future trends in Internet security. *Distributed Computing Systems, 1997*. Proceedings of the Sixth IEEE Computer Society Workshop 1997, 216–217.
3. Harris B, Hunt R. TCP/IP security threats and attack methods. *Computer Communications* February 1999.
4. Yongcheng L, Zhigang C, See-Mong T, Campbell RH. Security enhanced MPEG player. *Multimedia Software Development, 1996*. Proceedings International Workshop 1996, 169–175.
5. Zhao SongSheng, Lu XiCheng, Zhou XingMing. Dynamic quality of session control of real-time video multicast. *Intelligent Processing Systems, 1997. ICIPS '97*, 1737–1741.
6. Hwangjun Song, Jongwon K, Jay Kuo C-C. Real-time encoding frame rate control for H.263+ video over

the Internet. *Signal Processing: Image Communication* 1999.
7. Bolot J-C, Turletti T. A rate control mechanism for packet video in the Internet. *INFOCOM '94. Networking for Global Communications*. 13th Proceedings IEEE, 1994, Vol. 3, 1216–1223.
8. Spanos GA, Maples TB. Security for real-time MPEG compressed video in distributed multimedia applications. *Computers and Communications, 1996*. Conference Proceedings of the 1996 IEEE Fifteenth Annual International, Phoenix, 72–78.
9. Baukari N, Aljane A. Security and auditing of VPN services in distributed and networked environments. *Proceedings of Third International Workshop* 1996, 132–138.
10. Chen Y-W, Wu J-LC. A heuristic approach of bandwidth management for video sources in ATM networks. *International Journal of Network Management* 2000; **10**: 41–49.
11. Pao I-M, Sun M-T. A rate-control scheme for streaming video encoding. *Signals, Systems & Computers, 1998*. Conference Record of the Thirty-Second Asilomar Conference, Vol. 2, 1616–1620.
12. Gregory W, Cermak E, Tweedy P. Subjective evaluation of MPEG-2 video with and without B frames. *SPIE Conference on Multimedia Systems and Applications* 1999.
13. Kostas TJ, Borella MS, Sidhu I, Schuster GM, Grabiec J, Mahler J. Real-time voice over packet-switched networks. *IEEE Network* Jan./Feb. 1998.
14. Song H, Kim J, Jay Kuo C-C. Real-time encoding frame rate control for H.263+ video over the Internet. *Signal Processing: Image Communication* 1999.
15. Yeung A, Liew SC. Multiplexing video traffic using frame-skipping aggregation technique. *Image Processing 1997*. Proceedings International Conference, Vol. 1, 1997, 334–337.
16. Hwangjun S, Kuo C-CJ. Rate control of H.263 for low bit rate visual communication. *Signals, Systems & Computers 1997*. Conference Record of the Thirty-First Asilomar Conference, Vol. 1, 1998, 377–381.
17. Hwang JN, Wu T-D, Lin CW. Dynamic frame-skipping in video transcoding. *Multimedia-Signal Processing 1998*, IEEE Second Workshop. ■

