Performance analysis of secured video transmission over 3G networks based on H.264/AVC

Varalakshmi Kumar*

Department of Electronics and Communication Engineering,
Sri Manakula Vinayagar Engineering College,
Madagadipet, Puducherry, 605107, India
E-mail: varalakshmi_1@yahoo.co.in
*Corresponding author

Gnanou Florence Sudha

Department of Electronics and Communication Engineering,
Pondicherry Engineering College,
Pillaichavady, Puducherry, 605014, India
E-mail: gfsudha@ pec.edu

Ilamparuthi Paruthi

MBit Wireless Technologies,
Shivani Block I, GF No. 40, East Coast Road,
Thiruvanmiyur, Chennai, 600041, India
E-mail: Ilamparuthiparuthi@gmail.com

Abstract: Video transmission in wireless environment is a challenging task calling for high-compression efficiency as well as network friendly design which is achieved by means of H.264/AVC. Video transmission for mobile terminals is a major application in 3G systems and is a key factor for its success. Wireless links are error-prone and unreliable and so video transmission over 3G networks experiences packet loss due to fading and interference causing substantial quality degradation to the video which may be avoided by appropriate error resilience features. Moreover, it is important that the video is transmitted in a secured manner in certain applications like video conferencing, medical applications etc. This is achieved with the help of suitable encryption techniques without affecting the quality of the decoded video. Encryption of video transmission in wireless environment with suitable error resilience techniques to achieve good quality is the main objective of this work. Performance analysis of secured video transmission over 3G networks is done through quality measurement such as PSNR, bit overhead analysis, encryption time measurement, key sensitivity analysis and security analysis.

Keywords: video security; selective encryption ; AES; 3G networks; error-resilience; video quality; PSNR; multimedia transmission; H.264/AVC video codec; performance analysis.

Reference to this paper should be made as follows: Kumar, V., Sudha, G.F. and Paruthi, I. (2013) 'Performance analysis of secured video transmission over 3G networks based on H.264/AVC', *Int. J. Multimedia Intelligence and Security*, Vol. 3, No. 1, pp.23–50.

Biographical notes: Varalakshmi Kumar is currently working as an Associate Professor in the Department of Electronics and Communication Engineering at Sri Manakula Vinayagar Engineering College affiliated to Pondicherry University, Puducherry, India. She has completed her BE in Electronics and Communication Engineering from Thiagarajar College of Engineering, Madurai, and MTech in Electronics and Communication Engineering from Pondicherry Engineering College, Puducherry, India. She is currently pursuing her PhD in Multimedia Security under the guidance of Dr. Gnanou Florence Sudha.

Gnanou Florence Sudha is currently working as an Associate Professor in the Department of Electronics and Communication Engineering in Pondicherry Engineering College affiliated to Pondicherry University, Puducherry, India. She has completed her BTech and MTech in Electronics and Communication Engineering in 1992 and 1994 in Pondicherry Engineering College. She completed her PhD with specialisation in image processing in 2005. She has 16 years experience in teaching. She has several journal and conference proceedings to her credit. She is a member of the Indian Society of Technical Education, Optical and Biomedical Society of India. Her areas of research and interest are image processing, biomedical signal processing and information security.

Ilamparuthi Paruthi completed his BTech in Electronics and Communication Engineering at Pondicherry Engineering College, Puducherry in 2011 and currently working as a Development Engineer at MBit Wireless Technologies, Chennai.

1 Introduction

With the explosive growth of the internet and dramatic increase in wireless access, there is a tremendous demand on multimedia delivery over wireless internet. Moreover, the rapid growth of processing power and network bandwidth has caused many multimedia applications to emerge in the recent past. As digital data can be easily copied and modified, concern about its protection has gained importance in recent times. Multimedia data requires either full encryption or selective encryption depending on the application requirements (Wiegand et al., 2003; Sullivan and Wiegand, 2005). Conventional data encryption scheme is difficult to be applied directly due to the large volume of data and real-time transmission requirements. Selective encryption has been proposed recently as an ideal candidate to achieve data security (Lookabaugh and Sicker, 2004). This work studied the state-of-the-art multimedia privacy protection technologies and demonstrated that selective encryption is an efficient solution to achieve content protection. Lian et al. (2006) proposed a partial video encryption scheme in which the intra-prediction mode, residue data and motion vectors are encrypted based on the Exp-Golomb coding and context-adaptive variable length coding (CAVLC).

Selective encryption of H.264/AVC at network abstraction layer (NAL) has been proposed by Li et al. (2008). Important NAL units (NALU), namely instantaneous decoding refresh picture, sequence parameter set and picture parameter set are encrypted with a stream cipher. The limitation of this scheme is that it is not format compliant and cannot be parsed even at frame level. In one of the recent works (Shahid et al., 2011),

selective encryption is performed in both the types of entropy coding module, CAVLC and context-adaptive binary arithmetic coding (CABAC) using Advanced Encryption Standard (AES) on a subset of code words/bin strings. But the drawback of this scheme is its computational complexity particularly when applied to CABAC which is used only in higher profiles. As one of the means to protect media content, multimedia content encryption has attracted more and more researchers and engineers. Lian (2009) gives a detailed description and analysis on multimedia content encryption which includes a brief history of multimedia encryption, performance requirements, encryption techniques, attacks and performance evaluations. The *Handbook of Research on Secure Multimedia Distribution* (Lian, 2009) provides a scientific treatment of the state-of-the-art techniques related to multimedia distribution. However, most of the previous works on multimedia selective encryption have been focused mainly on the application layer cryptography and efficient delivery of encrypted video in error-prone wireless channels have not been considered much in literature.

H.264 uses the latest innovations in the video compression technology to provide high quality video at low data rates, making it a prime candidate for bandwidth limited wireless environments such as the third generation (3G) cellular networks. Error rate is usually very high in wireless channel caused by multi-path fading, inter symbol interference and noise disturbances. To reduce the errors in wireless channel, effective error control is essential for robust video transmission. Stockhammer et al. (2003) provides an overview of the error-resilience tools and the transmission characteristics for wireless video applications. It is also shown that as the length of the packets increases, packet-loss probability also increases significantly. This work served as a basis for analysing H.264 in wireless environments. Wenger (2003) describes the use of H.264 coded video over best-effort IP networks, using RTP as the real-time transport protocol. Liu et al. (2005) analysed the error-resilience tools and its usability in 3G networks. This work showed that the use of FMO mode and intra block refreshing results in good performance in terms of bit rate.

Wiegand et al. (2005) presented an audio/video frame interleaving scheme in 3G wireless environments. Simulations have included audio into the interleaving framework and the experimental results demonstrate the superior performance of interleaving for typical link outage settings. Qu et al. (2003) investigated the combined use of passive error concealment together with FEC coding and periodic intra updating to improve the performance in the presence of bursty packet losses. A robust cross-layer architecture that relies on a data partitioning technique at the application layer and an appropriate priority mapping at the 802.11e medium access control (MAC) layer is described by Ksentini et al. (2006). This solution is complemented by the schemes proposed by Fallah et al. (2007), where a cross-layer design is presented that optimises the encoded packet size to improve the H.264 video transmission over 802.11 wireless local area networks (WLAN). Instead of doing the fragmentation in the MAC layer, packet sizes are adjusted by slicing the video frame at the application layer, achieving better performance for a given packet-loss condition.

For mobile environment, 3G systems have a bandwidth of up to 384 kbps and for stationary environment 3G systems have a bandwidth up to 2 Mbps (Roth et al., 2001). The packet-loss rate of 3G mobile system is very high and the packet-loss model of 3G mobile systems is described (Yang and Bourbakis, 2009). Wang et al. (2000) reviewed the different error resilience techniques needed for real-time video transmission over unreliable networks. In addition to error resilience tools, the work analyses the error

concealment techniques and schemes that require forward error correction (FEC) and feedback to protect the video transmission. Kim et al. (2005) demonstrated that adding redundant information significantly improves the error concealing abilities of the decoder by sampling the texture coefficients and grouping into several interleaved groups to avoid the burst error damage of all groups within a single block in wireless video transmission. Calafate and Malumbres (2003) analysed the error resilience features suitable for wireless ad-hoc networks and suggested to tune the encoder with packet loss to increase the quality of the video. Ogunfunmi and Huang (2005) presented an improved three-dimensional flexible macro-block ordering (3D FMO) and achieved better video transmission qualities over wireless networks.

The contribution of this research is multifold. In this work, different from the previous works, a secure multimedia communication framework is proposed, to guarantee security, reduce packet loss over 3G networks and improve the quality of decoded video. The rest of the paper is organised as follows. In Section 2, overview of H.264/AVC and 3G networks is presented. The proposed scheme is discussed in detail in Section 3. Simulation results are analysed in Section 4 and performance analysis is discussed in Section 5. Finally concluding remarks are given in Section 6.

2 Background of H.264 over 3G networks

Video transmission is a fundamental service of the emerging 3G systems. Three major service categories are identified for 3G applications.

1 conversational services for video telephony and video conferencing

2 video streaming services

3 video in multimedia messaging services (MMS).

In general, mobile devices are hand-held and constrained in processing power and storage capacity. In addition, the mobile environment is characterised by harsh transmission conditions in terms of fading and multi-user interference which results in packet loss. Retransmission is used in delay insensitive applications like MMS and delivered error-free to the mobile user. In contrast, conversational and streaming services with real-time delay and jitter constraints allow for only a very limited number of retransmission.

In general, the available bandwidth and the bit-rate over the radio link are limited and the cost for the user is normally proportional to the number of bits transmitted over the radio link. Thus, low bit rates are likely to be typical and compression efficiency is the main requirement for a video coding standard to be successful in a mobile environment. This makes H.264/AVC a prime candidate for use in wireless systems because of its superior compression efficiency. Richardson (2010) gives a detailed analysis of the concepts, tools, benefits and disadvantages of the H.264 video compression standard and its significance to the broadcast, internet, consumer electronics, mobile security industries etc.

For efficient transmission in different environments, the seamless and easy integration of the coded video with all protocols and network architecture is important. The H.264/advanced video coding (AVC) structure are composed of two conceptual layers: the video coding layer (VCL) and the NAL. The VCL specifies an efficient

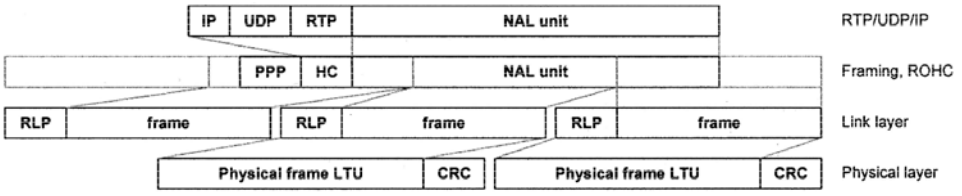
representation of the coded video signal. The NAL defines the interface between the video codec and the outside world. It operates on NALU, which supports packet switching networks, provides network friendliness and error robustness essential for real-time services. Each of the NALUs is encapsulated in a real time protocol (RTP) packet. For data transport either the user datagram protocol (UDP) or the transmission control protocol (TCP) is used. TCP cannot be used for real time communications due to its delay constraints. Thus, for real time data transfer UDP and IP are used both of which are unreliable protocols. UDP does not guarantee that data will reach the other end and IP only performs the best effort service. For this reason, there is a need to use RTP. RTP runs over UDP and with the help of features like sequence numbering, time stamp etc. RTP makes the transmission of video data feasible over unreliable transport and network layer protocols.

2.1 User plane protocols in 3G

3GPP and 3GPP2 define similar user plane protocols for UMTS and CDMA-2000 specifications between the mobile station (MS) and radio base station (BS). In CDMA-2000 a physical layer frame can be divided into a number of logical transmission units (LTU). The LTU is the smallest unit with a CRC to detect possible bit-errors. The size of an LTU (in bytes) is fixed for a session. An RLC-PDU in UMTS is of fixed length and is determined when the bearer is setup, for example 80 octets. For both it can be assumed, that one link layer frame is packed into one physical layer logical unit (LTU).

An RTP/UDP/IP packet is packed into one PDCP/PPP packet which becomes an RLC-SDU. Video packets are by nature of varying length, so RLC-SDUs will be of varying length and RLC-SDU which are not received correctly shall be discarded. It becomes important to select appropriate-sized RLC-SDUs (IP-packets) that match the channel conditions. The RLC/RLP layer can perform re-transmissions. The re-transmission scheme may be set up with different levels of persistency, such that the maximum number of re-transmission attempts (NAK rounds) is specified for a connection, which puts an upper limit to the variable RLC/RLP delay that re-transmission can cause. The RLC/RLP layer maintains in-order delivery, so usage of re-transmission can introduce high delay jitter. It is likely that conversational bearers are realised without RLC re-transmissions. However, for streaming applications the transfer delay requirements are typically in a range where RLC/RLP re-transmission is used. Figure 1 shows the packetization of application packets in the protocol stack.

Figure 1 Packetization through the user plane protocol stack



In order to be able to re-assemble the SDUs at the receiving side, the SDU boundaries must be signalled. In UMTS this is done by means of a length indicator in the RLC-PDU header. An RLC-PDU always starts with a sequence number (minimum size is 7 bits) and is followed by a 1-bit indicator which specifies if the next octet is a length indicator or

payload data. In CDMA-2000 the application packets are framed at the PPP layer using the PPP protocol and if applicable the RTP/UDP/IP header is compressed. Packetization overhead in UMTS protocol stack is:

- overhead per RLC PDU = RLC frame header (4 bytes) = 4 bytes
- overhead per PDCP PDU = PDCP packet header (1 byte) + length info (1 byte) + compressed RTP/UDP/IP header (3 bytes) = 5 bytes.

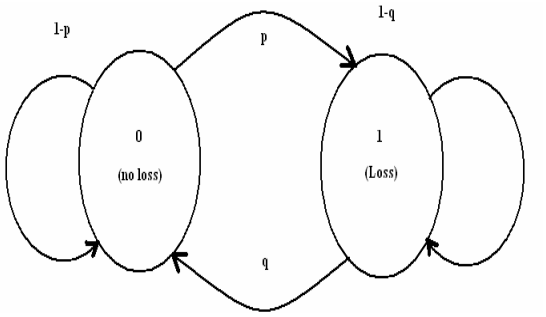
Packetization overhead in CDMA-2000 protocol stack:

- overhead per RLC PDU = RLC frame header (2 bytes) = 2 bytes
- overhead per PPP PDU = PPP packet header (3 bytes) + octet-synchronous HDLC stuffing (~2% of (payload + compressed RTP/UDP/IP header)) + compressed RTP/UDP/IP header (3 bytes).

2.2 Error resilience features of H.264/ AVC suitable for 3G services

In real-time multimedia applications, packet loss results from two situations: network loss and late loss. In network loss, packet is lost during transmission and cannot reach the destination. In late loss, the arriving packet is discarded because it arrives too late and these two types of packet loss make no difference to end users in terms of perceptual quality degradation. Packet loss behaviour can be approximated by first order Markov model also known as Gilbert model as shown in Figure 2, Markov Model is a powerful tool in describing the relationship between events that are temporally related with each other. In a first order General Markov Model, there are 2^1 = 2 possible states and 4 (2 * 2) possible transitions. Each state corresponds to two transitions: the transition to 0 and the transition to 1. State 0 means the packet reaches the receiver in time, and State 1 means packet loss. In the Gilbert Model, p is the probability that the next packet is lost provided that the previous one has arrived successfully, and q is the probability of the opposite case. 1-p is the probability that a delivered packet is followed by another delivered packet while 1-q is the probability that a lost packet (Figure 2).

Figure 2 Packet loss model



From the definition, π_0 and π_1 which are the state probabilities for 0 and 1, can be computed. In the Gilbert model π_0 and π_1 also represent the mean arrival and loss probability, respectively.

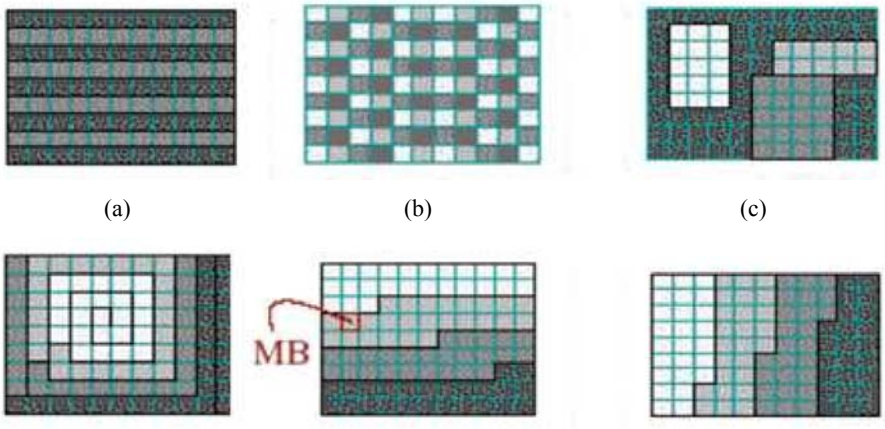
$$\pi_0 = q / p + q \qquad \pi_1 = p / p + q$$

Due to high packet loss rate in wireless environment, error resilience for 3G is necessary. The error resilient feature suitable for the transmission of video over 3G networks is discussed next using Baseline pro\file of JM encoder.

2.2.1 Flexible macro-block ordering

FMO techniques are based on data interleaving to decrease the influence of consecutive packet loss caused by burst errors. Using flexible macroblock ordering (FMO), each macro block (MB) in the frame can be assigned a slice group in a smart ordering way enabling easier reconstruction of missing blocks in the decoder side. This mechanism does MB rearrangement so that burst errors are distributed throughout the frame (Figure 3).

Figure 3 Standard FMO types, (a) interleave (b) dispersed (c) foreground with left-over (d) box-out (e) raster scan (f) wipe left (see online version for colours)



Scattering the error spatially within the video frame increases the probability that the neighbourhood of a corrupted MB is properly received, and therefore allows more effective error concealment by the decoder.

2.2.2 Intra block refreshing

H.264/AVC uses intra block refreshing to stop error propagation and drift caused by RTP packet loss. For real-time conversational video services, I frames are not inserted due to increased bit-rate and delay. Hence, intra block refreshing is very important to remove artefacts caused by inter frame prediction.

2.2.3 Error concealment

Error concealment utilises the received data to recover the lost data and commonly the spatial information and temporal information are used in the error recovery process. Error concealment is important in error prone environment, especially for wireless networks. The decoder plays a fundamental role in error resilience since it is responsible for error

concealment tasks. It maintains a status map for MBs to indicate if a particular MB has been correctly received, lost or already concealed, for each frame being decoded.

2.2.4 Feedback channel

Feedback channel can be efficiently used for packet retransmission and improve the performance of the system. The receiver can send ACK messages or NAK messages to the sender to report whether the RTP packets have been received or not. In the case of conversational applications, the encoder can select the corrected-received frames as reference frames. Feedback channel can efficiently improve the performance of the system. But packet retransmission may bring time delay to the decoder and it is restricted in 3G conversational video services.

2.2.5 Data partition and FEC

Because data partition is not supported by H.264/AVC baseline profile, unequal error protection (UEP) is also not supported in 3G video services. Moreover, packets with bit error are discarded in H.264/AVC mobile services and so FEC is also not suitable for 3G video services.

2.2.6 Redundant slices

Redundant slices (RS) is an error robustness feature which allows the encoder to send an extra representation of a picture region at lower fidelity that can be used if the primary representation is corrupted or lost. RS allow the insertion of primary slices and one or more additional secondary slices belonging to the original picture in the same bit stream. If a primary slice is affected by errors, it can be replaced by an error-free redundant one; otherwise the RS are discarded. As it introduces delay, it is not used in conversational and streaming services, which is the focus of this work.

3 Proposed work

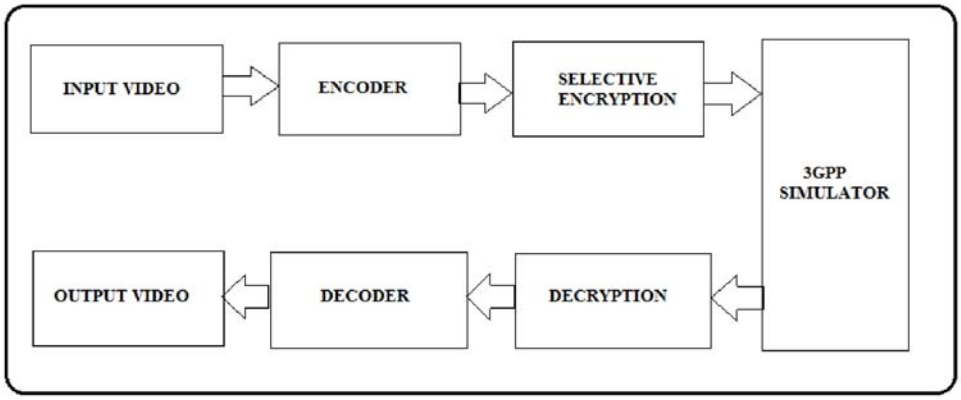
The main objective of the proposed work is to securely transmit the video over 3G networks. As encryption is combined with compression, high security strength can be achieved at a relatively low computation cost. The encryption scheme designed should offer high security and provide resistance against various attacks such as the ciphertext attack and plaintext attack. In addition the encryption time should be low which is usually set to be less than 5%. The ultimate goal of multimedia compression is to reduce the size of the bit stream to the maximum extent possible and the encryption scheme proposed should not violate this. Achieving high security at the expense of sacrificing the compression ratio is not desirable. Moreover the increase in the final bit stream size due to encryption should not be higher than 5% of the original coded bit stream.

Secured transmission is achieved by means of selective encryption applied to I frames and motion vectors. A cryptographic key is used to control the encryption. At the decoder, authorised users will then use the same key to reconstruct the original video. AES is used for encryption of I frames and P frames. AES is a symmetric-block cipher, has a complex structure and is an approved standard for a wide range of applications. The

cipher takes a plaintext block size of 128 bits and the key length chosen is also 128 bits. The encoded and the encrypted bit stream are then transmitted over 3G networks. 3G networks are characterised by harsh transmission conditions in terms of fading and multi-user interference, which results in packet loss.

The framework of the proposed scheme is shown in Figure 4. The framework can be formulated as a transmission quality optimisation problem with security and complexity constraints. Quality is measured in terms of PSNR, security in terms of bits selected for encryption and complexity in terms of encoding time. The number of bits selected for encryption is to be minimum by adjusting the error resilience features such as FMO, retransmission, data partition etc. so as to reduce the encoding time to attain maximum quality.

Figure 4 Framework of the proposed scheme



In general, security strength of an encryption scheme is decided by the total number of encrypted bits in each video frame as well as the key strength. Typically, encryption using a stronger block encryption algorithm with a longer key length provides a higher level of security. Encryption cost can be reduced by reducing algorithm complexity or key length. But this may result in security compromise. The encryption cost is lowered by reducing the total encryption workload, rather than reducing key length or algorithm complexity. In this work, AES is used to provide security and the cost is reduced by selectively encrypting the information bits. Furthermore, the selection of minimal encryption bits is restrained by the total encryption bits available, which is set by the user.

The expected video quality after decryption and decoding is maximised as shown in equation (1)

$$\underset{i\in 0,1\ldots N-1}{\{BL_i, ET_i\}} = \arg.\max\{Q_i\} \tag{1}$$

where

BL_i represents the encrypted bit length of i^{th} frame

ET_i denotes encoding and encrypting time of i^{th} frame

Q_i quality of i^{th} frame

N number of frames to be encoded.

subject to encryption bit constraint as given by equation (2)

$$\sum_{i\in 0,1\ldots N-1} BL_i \le TB_i \tag{2}$$

and minimum security constraint as given by equation (3)

$$\sum_{i\in 0,1\ldots N-1} BL_i \ge BL_{\min} \tag{3}$$

where

TB_i represents the total bit allocated for encryption of i^{th} frame

$BL_{\min}$ represents the minimum security requirement

with complexity constraint as shown in equation (4)

$$\sum_{i\in 0,1\ldots N-1} ET_i \le ET_{\max} \tag{4}$$

and $ET_{\max}$ is the maximum time allocated for encoding of i^{th} frame. In the proposed scheme, the minimum security requirement ($BL_{\min}$) is set by the user and the ideal encryption length is acquired by ceiling the minimal security requirement bits to the lowest exponential value of a binary two as shown in equation (5).

$$BL_i = 2\left\lceil \log_2 \left(BL_{\min}\right)\right\rceil^2 \tag{5}$$

where

$$\left\lceil \log_2 \left(BL_{\min}\right)\right\rceil^2$$

be greater than $BL_{\min}$ and less than $TB_{i,}$ the total bit allocated for encryption of i^{th} frame.

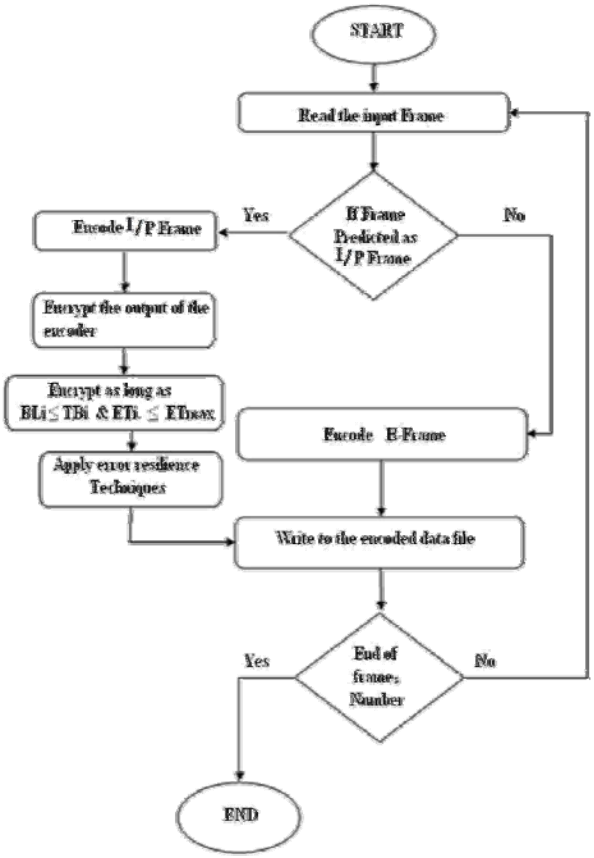
Encryption of I/P frame is performed during encoding and the number of bits encrypted is to be less than TB_i and in addition the encoding time has to be less than $ET_{\max}$. In the proposed work, this value is set to be approximately 5% more than the encoding time alone. This is to indicate that the time taken for joint encryption and encoding is not to exceed 5% of the time taken for encoding alone. The algorithm and flowchart for I/P frame encryption is given in Figure 5.

Algorithm for I/P frame encryption

- encode the video using JM encoder.
- assign parameters such as input video, number of frames to be encoded, format of output, period of I/P frames etc.
- I frame is identified by the hexadecimal value 0 × 65
- P frame is identified by the hexadecimal value 0 × 41
- I/P frame is fragmented into blocks of 128 bits
- each block is given as input to the AES algorithm

- the key K_u is obtained from the user
- the ciphertext obtained for each 128 bit block replaces the original 128 bit block of the I frame
- update the number of encrypted bits
- if the number of encrypted bits is less than the encryption bit budget, continue processing the next MB
- the above step continues as long as $BL_i \leq TB_i$ and $ET_i. \leq ET_{\max}$
- repeat for other I/P frames in the encoded video
- enable error concealment
- enable FMO
- enable intra MB refresh
- disable feedback channel(conversational)/enable feedback (streaming)
- packetize and transmit through 3G networks.

Figure 5 Flowchart of I/P frame encryption



I frame of an encoded video contains most of the information. P frames use I frames as reference, hence the encryption of an I frame affects the corresponding P frames in the video. When an attack is made by the malicious user it is not possible for the attacker to decode the video, due to I frame encryption. For commercial TV broadcast, encryption of I frame alone is sufficient. But in the case of military applications where tight security is needed P frames have to be encrypted in addition to I frames. P frames are the inter predicted frames and coded by differences in the motion between two different frames. The P frames can be identified by a value 0 × 41 when the .264 file is viewed in a Hex Viewer. This value denotes the start of a P frame. When the P frame is encrypted the motion vectors in the P frames are changed, leading to irrelevant pictures and the flow of the video is affected by encryption of P frames. Figure 6 and Figure 7 portrays the effect of I frame encryption. Figure 8 and Figure 9 portrays the effect of combined I frame and P frame encryption.

The frame sequence is also viewed in Elecard stream analyser as shown in Figure 10(a). The tallest frame (purple) represents the I frame and the remaining frames denote the P frames. The encrypted I frame is denoted by colour change as shown in Figure 10(b).

Figure 6 Coastguard frame, (a) original (b) I frame encrypted (c) reconstructed (see online version for colours)

(a) (b) (c)

Figure 7 Foreman frame (a) original (b) I frame encrypted (c) reconstructed (see online version for colours)

(a) (b) (c)

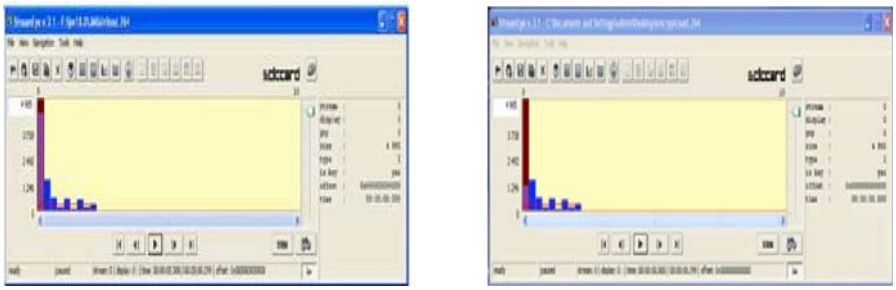
Figure 8 Coastguard frame, (a) original (b) I and P frame encrypted (c) reconstructed (see online version for colours)

(a) (b) (c)

Figure 9 Foreman frame, (a) original (b) I and P frame encrypted (c) reconstructed (see online version for colours)

(a) (b) (c)

Figure 10 (a) Snapshot of ten frames of hall sequence (IPPP..) (b) Snapshot of encrypted I frame alone (see online version for colours)



(a) (b)

The encrypted video is transmitted over 3G networks using various error resilience features. The received video is first decrypted and then decoded. The quality of this decoded video is measured in terms of PSNR value for different error resilience combinations and packet loss ratios simulated using 3GPP simulator.

4 Simulation results and analysis

The input video is first encoded using JM18.0 software. The encoded bit stream is selectively encrypted using AES and then packetized and transmitted over 3G networks using 3GPP2 simulator. 3GPP is an offline simulator used to simulate the characteristics of 3G wireless networks. Both video source and receiver are assumed to reside in a private operator's network, which consists of a fixed IP-based core network and a radio access network. A video streaming server is directly connected to the operator's core network or both video call terminations are inside the mobile network.

The following assumptions have been made for the simulation:

- off-line simulation
- in-order packet delivery
- no bit-errors or corrupted packets delivered to the application
- no UEP
- simulate delay jitter in case of RLC re-transmission or no delay jitter otherwise.

Table 1 defines the video sequences and bitrates used in experiments.

Table 1 Video sequences used

Sequence	*Picture size*	*Frame-rate (f/s)*	*To be used with physical layer rate (kbits/sec)*
Hall monitor	QCIF and CIF	15	64
Coastguard	QCIF and CIF	10	64
Foreman	QCIF and CIF	10	64

For simulating radio channel conditions bit-error patterns are used that were captured in different real or emulated mobile radio channels. The bit-error patterns are captured above the physical layer and below the RLC/RLP layer, such that in practice they act as the physical layer simulation. Table 2 lists the bit-error patterns used in the experiments for simulating radio channel conditions. Patterns 1and 2 are mostly suited for video streaming applications, where RLC re-transmissions can correct many of the frame losses. Patterns 3 to 6 are meant for video conversational applications.

Table 2 Bit error patterns

S. no.	*File name*	*Bit rate*	*Length*	*BER*	*RLC PDU size*	*Mobile speed*
1	18681.3	64 kbps	60 s	9.3e-3	640 bits	3 km/h
2	18681.4	64 kbps	60 s	2.9e-3	640 bits	3 km/h
3	WCDMA_64kb_3kph_5e-04.bin	64 kbps	180 s	5.1e-4	640 bits	3 km/h
4	WCDMA_64kb_50kph_2e-04.bin	64 kbps	180 s	1.7e-4	640 bits	50km/h
5	WCDMA_128kb_3kph_5e-04.bin	128 kbps	180 s	5.0e-4	640 bits	3 km/h
6	WCDMA_128kb_50kph_2e-04.bin	128 kbps	180 s	2.0e-4	640 bits	50km/h

Source: Roth et al.

The input sequence is encoded once and the encoded bit stream is run through the simulator ten times with different starting position of the bit-error pattern, that are specified for each bit-error file as shown in Table 3.

Table 3 Simulation starting positions for the different bit-error patterns

No.	*Simulation starting positions (in bytes)*
1	227,200, 259,840, 128,000, 348,800, 81,920, 79,360, 392,320, 56,960, 440,960, 458,880
2	227,200, 259,840, 128,000, 348,800, 81,920, 79,360, 392,320, 56,960, 440,960, 458,880
3	551,040, 251,840, 613,120, 609,920, 616,000, 57,600, 1,313,120, 1,248,640, 273,920, 1,366,400
4	551,040, 251,840, 613,120, 609,920, 616,000, 57,600, 1,313,120, 1,248,640, 273,920, 1,366,400
5	1,143,200, 410,240, 2,480,000, 561,920, 2,135,840, 1,954,560, 283,520,730,880, 698,720, 154,368
6	1,143,200, 410,240, 2,480,000, 561,920, 2,135,840, 1,954,560, 283,520, 730,880, 698,720, 154,368

Source: Roth et al.

The encoded sequence is of IPPP… format with QP = 32, profile IDC = 66, level IDC = 40. Packet loss is generated by off-line simulation software and when a frame cannot be recovered due to RTP packets loss, direct frame copy from previous frame is performed. The error resilience strategies adopted in this work is listed in Table 4.

Conversational service on 3G networks is a challenging task, because mobile terminals must have the real-time ability of both encoding and decoding with low computational complexity. Error concealment is the most important tool, because feedback channel and RTP packet retransmission are not allowed due to the implicit time delay. However, slice structure with FMO is used. Moreover, intra block refreshing is also supported to remove the error drift caused by the loss of inter-prediction MB. RS will not be supported, as it increases the bit rate.

Table 4 Error resilience strategies for mobile video services

Scheme	*Conversational applications*	*Streaming applications*
Error concealment	Yes	Yes
Redundant slices	No	No
FMO	Yes	Yes
Intra MB refresh	Yes	Yes
Feedback channel	No	Yes

Table 5 Simulation results – FMO types

FMO types	*PSNR (dB)*	*Bit rate (kb/s)*	*Encoding time (sec)*
Type 0 (interleave)	35.125	86.227	281.926
Type 1 (dispersed)	35.079	101.414	282.031
Type 2 (foreground with leftover)	35.116	80.434	282.389
Type 3 (box-out)	35.117	82.694	281.609
Type 4 (raster scan)	35.129	82.238	295.421
Type 5 (wipe)	35.074	82.570	281.769

Table 6(a) Experimental results of 3G conversational services without error resilience

Video sequence	*Frame rate (fps)*	*RTP packets*	*Lost packets*	*Bits transmitted*	*Bits in error*	*Frames transmitted*	*Frames in error*	*Encoded ave. Y PSNR*	*Decoded ave. Y PSNR*
Hall monitor	15	800	64	1,163,520	293	9,090	59	32. 348	23.182
Foreman	15	748	57	1,153,920	280	9,015	524	34.029	29.216
Coastguard	10	971	270	1,720,064	1,912	13,438	1,972	26.319	20.527

Notes: Video format: QCIF; total frames: 300; bit error pattern: 4

Table 6(b) Experimental results of 3G conversational services with error resilience

Video sequence	*Frame rate (fps)*	*RTP packets*	*Lost packets*	*Bits transmitted*	*Bits in error*	*Frames transmitted*	*Frames in error*	*Encoded ave. Y PSNR*	*Decoded ave. Y PSNR*
Hall monitor	15	800	5	1,163,520	68	9,090	24	35.124	28.490
Foreman	15	748	6	1,153,920	68	9,015	24	39.358	31.532
Coastguard	10	971	208	1,720,064	1,704	13,438	1,630	28.517	23.461

Notes: Video format: QCIF; total frames: 300; bit error pattern: 4

Table 7(a) Experimental results of video streaming services without error resilience

Video sequence	*Frame rate(fps)*	*RTP packets*	*Lost packets*	*Bits transmitted*	*Bits in error*	*Frames transmitted*	*Frames in error*	*Encoded ave. Y PSNR*	*Decoded ave. Y PSNR*
Hall monitor	15	800	78	1,163,520	326	9,090	65	30. 729	21.624
Foreman	15	748	192	1,153,920	13,152	9,015	1,213	35.914	28.522
Coastguard	10	971	245	1,720,064	18,538	13,438	1,785	28.261	23.173

Notes: Video format: QCIF; total frames: 300; bit error pattern: 1

Table 7(b) Experimental results of video streaming services with error resilience

Video sequence	*Frame rate(fps)*	*RTP packets*	*Lost packets*	*Bits transmitted*	*Bits in error*	*Frames transmitted*	*Frames in error*	*Encoded ave. Y PSNR*	*Decoded ave. Y PSNR*
Hall monitor	15	800	20	1,163,520	168	9,090	31	34.712	22.490
Foreman	15	748	130	1,153,920	11,965	9,015	930	39.644	32.652
Coastguard	10	971	167	1,720,064	16,032	13,438	1,226	32.779	26.741

Notes: Video format: QCIF; total frames: 300; bit error pattern: 1

For video streaming services, the bit stream is typically pre-encoded and transmitted on demand through unreliable transmission protocols making error resilience necessary. Besides, due to the tolerance of relatively long delay compared to conversational applications, feedback channel is used in addition to intra MB refreshing and FMO. Simulation results of the various FMO modes are shown in Table 5.

From Table 5 it is seen that type 2 is preferred over other schemes as it offers comparatively good PSNR at a low bit rate and similar encoding time. Hence type2 FMO mode is used for transmission. Experimental results of 3G conversational services and video streaming services with/without error resilience are shown in Table 6(a) and Table 6(b) and Table 7(a) and Table 7(b). Results in Table 6a and 6b show that conversational video transmission with these error resilience schemes can efficiently stop error propagation and error drift and obtain satisfactory results.

Error resilience schemes with simple FMO and intra block refreshing is preferred as it can make full use of the spatial correlation in the MBs for intra prediction and can efficiently reduce the requirement of bit rate for encrypted video. From Table 7(a) and Table 7(b) it is seen that RTP packet loss is high without error resilience and low with resilience as shown by reduced packet loss and improved PSNR value.

Analysis is done in 3G networks with error pattern 4. The test video is encoded and encrypted once and is tested with ten different starting positions specified for error pattern 4 as shown in Table 8.

Table 8 Runs for error pattern 4

S. no.	*Starting position*	*Total bits transmitted*	*Bits in error*	*RTP packets*	*Lost packets*	*Frames transmitted*	*Frames in error*
1	551,040	189,568	0	423	0	1,481	0
2	251,840	189,568	2	423	1	1,481	1
3	613,120	189,568	43	423	10	1,481	18
4	609,920	189,568	43	423	10	1,481	18
5	616,000	189,568	43	423	10	1,481	18
6	57,600	189,568	26	423	6	1,481	10
7	1,313,120	189,568	332	423	7	1,481	20
8	1,248,640	189,568	16	423	7	1,481	9
9	273,920	189,568	284	423	7	1,481	20
10	1,366,400	189,568	172	423	6	1,481	20

PSNR of the decoded video is calculated for each of the ten runs and the average PSNR plus the best and worst cases of the ten runs are shown. This method is used to show the variation of the PSNR depending on the position of the losses. In 3G networks any constant bit rate is guaranteed by the network (<2,048 kbit/s). If the video stream bit rate is always equal or lower than the guaranteed bit rate, the application does not have to perform rate adaptation. For a given maximum application SDU size an SDU loss ratio can be guaranteed. This means, that the error robustness of a video packet stream should be tuned to a given packet loss ratio, which also means, that the video packet size influences the packet loss rate. Packet sizes are in turn adjusted by slicing the video frame at the application layer, achieving better performance for a given packet-loss condition.

Hence, the slice size plays an important role in the quality of the video. PSNR and bit rate obtained for different slice sizes are shown in Table 9.

Table 9 Analysis of slice size

S. no.	*Slice size (bytes)*	*Foreman*		*Hall*		*Coastguard*	
		Bit rate (kb/s)	*PSNR (dB)*	*Bit rate (kb/s)*	*PSNR (dB)*	*Bit rate (kb/s)*	*PSNR (dB)*
1	50	73.420	33.960	81.21	31.485	75.31	31.779
2	100	69.176	39.678	78.43	35.006	73.64	32.115
3	200	67.976	39.644	73.90	35.766	73.51	32.779
4	400	67.624	39.648	70.03	35.967	73.37	32.934
5	600	67.560	39.653	69.25	35. 236	73.21	32.779
6	800	67.168	39.645	68.69	35. 418	73.20	33.009
7	1,000	66.648	39.690	68.02	35.521	73.14	32.040

Figure 11 Slice size vs. bit rate (see online version for colours)

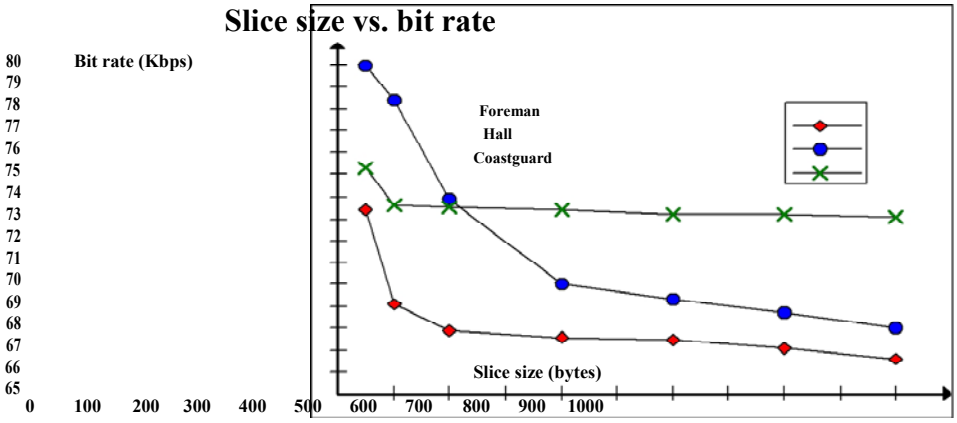
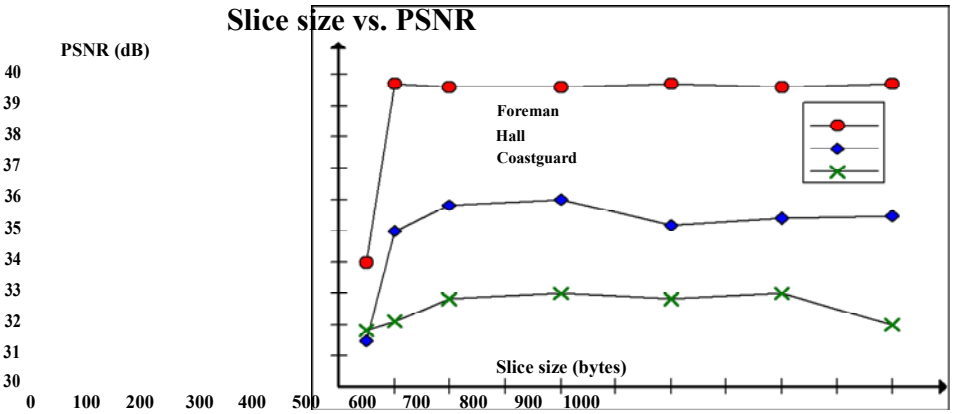


Figure 12 Slice size vs. PSNR (see online version for colours)



From Table 9 it is seen that as the slice size increases, bit rate decreases as illustrated in Figure 11. To transmit the video over 3G networks, the video bit rate and the bit rate due to the overhead associated with each packet should match the channel capacity. As the video stream bit rate is greater than the guaranteed bit rate (64 kbps), the application has to perform rate adaptation.

Figure 12 shows the PSNR value for various slice sizes. PSNR is initially high and later becomes nearly constant from a slice size of 200.

Table 10 PSNR Comparison without and with encryption

Test sequence	*Without encryption PSNR (dB)*	*CAVLC encryption PSNR (dB) (Shahid et al.)*	*Proposed I frame encryption PSNR (dB)*	*Proposed I + P frame encryption PSNR (dB)*
Foreman (QCIF)				
QP 24	38.37	18.62.	8.24	8.02
28	36.62	19.57	8.91	8.43
32	35.91	20.80	9.32	8.96
36	34.57	21.24	9.57	9.21
Silent (QCIF)				
QP 24	37.48	19.18	7.93	7.13
28	35.31	20.47	8.07	7.56
32	34.73	22.84	8.92	8.17
36	32.86	22.63	9.11	8.49
Hall				
Monitor (QCIF)				
QP 24	34.79	15.92	7.33	6.82
28	33.70	17.03	7.84	7.11
32	32.81	18.74	8.02	7.34
36	31.92	19.62	8.26	7. 67
Mobile (CIF)				
QP 24	39.68	17.75	8.01	7.34
28	38.07	19.21	8.28	7.85
32	37.81	20.84	8.95	8.24
36	36.34	21.53	9.04	8.61
Coastguard (CIF)				
QP 24	32.84	15.99	7.11	6.43
28	30.59	16.37	7.64	6.93
32	29.41	17.03	7.83	7.15
36	28.04	17.92	8.07	7.52
Stefan (CIF)				
QP 24	40.61	20.34	8.21	7.57
28	38.44	20.71	8.39	7.72
32	36.82	21.08	8.60	7.96
36	35.01	21.72	8.94	8.23

Table 11 Bit length (overhead bit) comparison of the encryption schemes

Test sequence		*Without encryption (bits)*	*CAVLC encryption (Shahid et al.) (bits)*	*Proposed I frame encryption (bits)*	*Proposed I + P frame encryption (bits)*
Foreman (QCIF)					
QP	24	41,328	41,432	41,445	41,505
	28	41,304	41,332	41,341	41,440
	32	41,298	41,321	41,304	41,415
	36	41,284	41,296	41,285	41,350
Silent (QCIF)					
QP	24	43,725	43,861	43,902	43,985
	28	43,712	43,734	43,741	43,825
	32	43,692	43,706	43,699	43,815
	36	43,674	43,637	43,680	43,720
Hall					
Monitor (QCIF)					
QP	24	45,341	45,397	45,402	45,450
	28	45,315	45,384	45,382	45,502
	32	45,294	45,375	45,346	45,475
	36	45,279	45,304	45,321	45,590
Mobile (CIF)					
QP 24		64,579	64,596	64,583	64,610
	28	64,558	64,573	64,571	64,615
	32	64,535	64,564	64,558	64,605
	36	64,502	64,670	64,669	64,620
Coastguard (CIF)					
QP 24		42,748	42,869	42,875	42,905
	28	42,721	42,853	42,864	42,915
	32	42,703	42,821	42,805	42,952
	36	42,691	42,789	42,775	42,818
Stefan (CIF)					
QP	24	46,558	46,717	46,795	46,817
	28	46,527	46,604	46,673	46,870
	32	46,501	46,571	46,562	46,570
	36	46,484	46,507	46,495	46,552

5 Performance analysis of the encryption scheme

Performance analysis of the proposed scheme is based on PSNR measurement, key sensitivity analysis and security analysis. To examine the effectiveness of the proposed algorithm in different experimental conditions, six benchmark video sequences [three in

quarter common intermediate format (QCIF) format (176 × 144) and three in common intermediate format (CIF) format (352 × 288)] was tested with IPPP.. frame structure and CAVLC entropy coding. All the video sequences are encoded with rate distortion optimisation (RDO) enabled.

Table 12 Encryption efficiency comparison

Test	*Encryption/compression (%)*			*Decryption/decompression(%)*		
Video sequence	*CAVLC scheme. (Shahid et al.)*	*Proposed I frame encryption*	*Proposed I + P frame encryption*	*CAVLC scheme (Shahid et al.)*	*Proposed I frame encryption*	*Proposed I + P frame encryption*
Foreman (QCIF)						
24	0.80	0.81	0.83	4.4	4.5	4.7
28	0.80	0.80	0.83	4.2	4.3	4.5
32	0.79	0.78	0.82	4.1	4.3	4.4
36	0.77	0.77	0.81	4.1	4.2	4.4
Silent (QCIF)						
24	0.73	0.74	0.76	4.5	4.7	4.9
28	0.73	0.73	0.77	4.4	4.6	4.8
32	0.71	0.73	0.76	4.4	4.5	4.8
36	0.70	0.72	0.75	4.3	4.3	4.6
Hall Monitor (QCIF)						
24	0.70	0.71	0.74	5.4	5.5	5.7
28	0.70	0.70	0.74	5.3	5.3	5.6
32	0.67	0.69	0.73	5.1	5.2	5.4
36	0.66	0.67	0.73	5.0	5.0	5.3
Mobile (CIF)						
24	0. 90	0. 92	0. 95	6.7	6.8	7.2
28	0. 89	0. 91	0. 94	6.5	6.7	7.1
32	0. 89	0. 91	0. 95	6.4	6.5	6.8
36	0. 88	0. 90	0. 93	6.4	6.4	6.7
Coastguard (CIF)						
24	0. 82	0. 83	0. 86	6.0	6.1	6.4
28	0. 82	0. 82	0. 86	5.9	5.8	5. 9
32	0. 81	0. 81	0. 84	5.7	5.7	5.9
36	0. 80	0. 81	0. 85	5.4	5.5	5.8
Stefan (CIF)						
24	1.2	1.3	1.6	7.1	7.2	7.4
28	1.1	1.2	1.5	7.0	7.1	7.3
32	1.0	1.2	1.4	6.9	6.8	7.1
36	1.0	1.1	1.4	6.7	6.7	6.9

5.1 Perceptual quality of video

PSNR is the primary means of measuring visual distortion and is the most widely accepted objective measure of visual distortion. Table 10 compares the PSNR values of the different video sequences, without and with encryption. To examine the performance at different bit rates, the experiments were performed with four quantisation parameter (QP) values, namely 24, 28, 32 and 36. It is seen that with increase in QP, the quality of the encrypted video increases. It is because of the fact that with higher QP, there are lesser non-zero coefficients. But, irrespective of the QP value, PSNR value of the encrypted video is always lower than that of unencrypted video. The proposed encryption scheme is compared with the context adaptive variable length encoding (CAVLC) scheme discussed by Shahid et al. both of which are based on variable length coding for encryption and uses the Baseline profile of H.264 for encoding as shown in Table 10.

Table 10 shows that the proposed scheme has scrambled the video better than the CAVLC scheme making it more unintelligible.

5.2 Bit overhead analysis

Table 11 compares the encrypted bit length of the proposed scheme with the CAVLC encryption scheme. The percent increase in the bits of the proposed scheme due to encryption is very less as compared to the original video, offering comparable performance with the previous scheme, and showing negligible overhead making it suitable for real time applications.

5.3 Encryption efficiency

Because real-time transmission is often required by multimedia applications, encryption algorithms should be efficient to minimise delay during transmission. As encryption is frequently used with compression, the comparative ratio between encryption and compression is used as a measure of encryption speed as shown in Table 12. It is seen that the proposed scheme is comparable in encryption speed with the CAVLC scheme for different QP values and the time overhead of the proposed scheme is negligible due to encryption /decryption.

5.4 Security analysis

For multimedia data, an encryption algorithm is said to be computationally secure if the cost to break the encryption by a cryptanalyst needs more investment than buying the key itself. There are multiple ways of attacking the encrypted video. The security of the proposed algorithm is checked against most common ways of attacking the video like ciphertext-only attack and known-plaintext attack.

5.4.1 Ciphertext-only attack

In this attack scenario, the attacker can only have access to the encrypted bit stream. Since the information available to the attacker is very limited, a very commonly used approach is the brute-force attack, the complexity of which is related to the key space. Since the only private information of the proposed system is the seed used in the stream

cipher and it is of length 128 bits, then the key space is 2^ 128, which can ensure satisfactory level of security.

5.4.2 Chosen ciphertext attack

In the known-plaintext attack, unauthorised user has both original and the corresponding encrypted values. For a cipher, an important property is that the same key stream should not be used more than once. Even if an attacker somehow finds a key stream, it is useless since it will not be used again. This essentially precludes the chosen-ciphertext attack. Prior to communication between the encoder and the decoder, a common seed, has to be exchanged between the two. The security of the stream cipher heavily relies on the fact that the same key stream will not be used for more than once. Therefore, if the same seed is used to generate the key stream used in different sessions, the security will be greatly degraded. Hence a new seed is used for each session.

5.5 Key sensitivity analysis

The encryption scheme exhibits high key sensitivity, which indicates that even slight difference in the keys can cause great difference in the decrypted video. Figure 13 shows the key sensitivity analysis. Figure 13(a) shows the original frame of Stefan sequence, Figure 13(b), Figure 13(c) and Figure 13(d) shows the encrypted frame, frame decrypted with wrong key and the frame decrypted with correct key respectively.

Figure 13 Key sensitivity analysis, (a) original frame (b) encrypted frame (key – 11001001) (c) decrypted with wrong key (11001010) (d) decrypted with correct key (11001001) (see online version for colours)

(a) (b) (c) (d)

This shows that the encryption scheme is of high key sensitivity, which makes statistical attacks very difficult.

6 Conclusions

The proposed work performed selective encryption of H.264 video, analysed the various error resilience tools suitable for transmission of encrypted and unencrypted video over 3G networks and identified the appropriate scheme to achieve good quality video. Experiments showed that when RTP packets are lost, the PSNR value falls rapidly and has a permanent adverse effect on the following frames. Use of appropriate error resilience strategy can prevent falling of PSNR value and efficiently remove error drift to obtain satisfactory PSNR value. The proposed encryption scheme is compared with the

CAVLC scheme by Shahid et. al., and the comparison results demonstrates that the proposed method scrambles the video better than the previous scheme, while producing only small additional overhead. In addition security analysis and key sensitivity analysis was performed to check the strength of the encryption scheme.

References

Calafate C.M. and Malumbres, M.P. (2003) 'Testing the H.264 error-resilience on wireless ad-hoc networks', *EURASIP Video/Image Processing and Multimedia Communications Conference*.

Fallah, Y.P. et al. (2007) 'A cross layer optimization mechanism to improve H.264 video transmission over WLANs', *Proc. of Consumer, Communications and Networking Conference (CCNC) 2007*, January, pp.875–879.

Kim, M. et al. (2005) 'Error resilient texture coding scheme for wireless video transmission based on coefficient sampling and interleaving', *SPIE Optical Engineering*, December.

Ksentini, A. et al. (2006) 'Toward an improvement of H.264 video transmission over IEEE 802.11e through a cross-layer architecture', *IEEE Communications Magazine*, January, Vol. 44, No. 1, pp.107–114.

Li, C. et al. (2008) 'NAL level encryption for scalable video coding', *Proc. PCM*, No. 5353., pp.496–505.

Lian, S. (2009) *Multimedia Content Encryption. Techniques and Applications*, CRC Press, Boca Raton, FL.

Lian, S. et al. (2006) 'Secure advanced video coding based on selective encryption algorithms', *IEEE Trans. Consumer Electronics*, May, Vol. 52, No. 2, pp.621–629.

Liu, L. et al. (2005) 'Error resilience schemes of H.264/AVC for 3G conversational video services', *Proceedings of the Fifth International Conference on Computer and Information Technology (CIT'05)*, September, pp.657–661.

Lookabaugh, T. and Sicker, D.C. (2004) 'Selective encryption for consumer applications', *IEEE Communication Mag.*, May, Vol. 42, No. 5, pp.124–129.

Ogunfunmi, T. and Huang, W.C. (2005) 'A new flexible macro block ordering with 3D MBA map for H.264/AVC', *IEEE ISCAS*.

Qu, Q. et al. (2003) 'Robust H.264 video coding and transmission over bursty packet-loss wireless networks', *IEEE Vehicular Technology Conference, VTC 2003*, Fall, October, Vol. 5, pp.3395–3399.

Richardson, I.E. (2010) *The H.264 Advanced Video Compression Standard*, John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK.

Roth, G. et al. (2001) *Common Test Conditions for RTP/IP over 3GPP/3GPP2*, ITU-T SG16 Doc. VCEGM77, Austin, TX.

Shahid, Z. et al. (2011) 'Fast protection of H.264/AVC by selective encryption of CAVLC and CABAC for I and P frames', *IEEE Transactions on Circuits and Systems for Video Technology*, May, Vol. 21, No. 5, pp.565–576.

Stockhammer, T., Hannuksela, M.M. and Wiegand, T. (2003) 'H.264/AVC in wireless environments', *IEEE Transactions on Circuits and Systems for Video Technology*, July, Vol. 13, No. 7, pp.657–673.

Sullivan, G.J. and Wiegand, T. (2005) 'Video compression – from concepts to the H.264/AVC standard', *Proc. IEEE*, January, Vol. 93, No. 1, pp.18–31.

Wang, Y. et al. (2000) 'Error resilient video coding techniques', *IEEE Signal Processing Magazine*, July, pp.61–82.

Wenger, S. (2003) 'H.264/AVC over IP', *IEEE Trans. on Circuits and Systems for Video Tech.*, July, Vol. 13, No. 7, pp.645–656.

Wiegand, T. et al. (2003) 'Overview of the H.264/AVC video coding standard', *IEEE Trans. Circuits Syst. Video Technol.*, July, Vol. 13, No. 7, pp.560–576.

Wiegand, T. et al. (2005) 'H.264/AVC interleaving for 3G wireless video streaming', *Multimedia and Expo., ICME 2005. IEEE International Conference*, 6–8 July.

Yang, M. and Bourbakis, N.G. (2009) 'Packet loss recovery methodology for video streaming over IP networks', *IEEE Transactions on Broadcasting*, June, Vol. 55, No. 2, pp.190–201.