International Journal of Electrical, Electronic and Communication Sciences
ISSN: 2517-9438
Vol:5, No:7, 2011

Selective Encryption using ISMACryp in Real Time Video Streaming of H.264/AVC for DVB-H Application

Jay M. Joshi, Upena D. Dalal

***Abstract*—**Multimedia information availability has increased dramatically with the advent of video broadcasting on handheld devices. But with this availability comes problems of maintaining the security of information that is displayed in public. ISMA Encryption and Authentication (ISMACryp) is one of the chosen technologies for service protection in DVB-H (Digital Video Broadcasting-Handheld), the TV system for portable handheld devices. The ISMACryp is encoded with H.264/AVC (advanced video coding), while leaving all structural data as it is. Two modes of ISMACryp are available; the CTR mode (Counter type) and CBC mode (Cipher Block Chaining) mode. Both modes of ISMACryp are based on 128-bit AES algorithm. AES algorithms are more complex and require larger time for execution which is not suitable for real time application like live TV. The proposed system aims to gain a deep understanding of video data security on multimedia technologies and to provide security for real time video applications using selective encryption for H.264/AVC. Five level of security proposed in this paper based on the content of NAL unit in Baseline Constrain profile of H.264/AVC. The selective encryption in different levels provides encryption of intra-prediction mode, residue data, inter-prediction mode or motion vectors only. Experimental results shown in this paper described that fifth level which is ISMACryp provide higher level of security with more encryption time and the one level provide lower level of security by encrypting only motion vectors with lower execution time without compromise on compression and quality of visual content. This encryption scheme with compression process with low cost, and keeps the file format unchanged with some direct operations supported. Simulation was being carried out in Matlab.

***Keywords*—**AES-128, CAVLC, H.264, ISMACryp

I. Introduction

ISMA Encryption and Authentication is one of the three chosen technologies for service protection in DVB-H (Digital Video Broadcasting - Handheld), the TV system for portable handheld devices [1]. The Internet Streaming Media Alliance (ISMA) standardized the first version of ISMA authentication and encryption in 2003, usually abbreviated as ISMACryp 1.0 [2]. The ISMA 1.0 is encoded with MPEG-4 video and MPEG-4 AAC (advanced audio coding), while leaving all structural data as it is. After the ISMA 2.0 specification, which is extended the prior version with profiles for the H.264/AVC video and HE-AAC audio codec [3]. ISMACryp works on top of the application layer and is usually inserted as a block just before the decoder. ISMACryp provides complete end-to-end protection with access units being protected by the content creator and being decrypted only right before decoding [3].

The ISMACryp algorithm for DVB-H makes the use of stream cipher that are used to encrypt/decrypt the frame encoded/decoded video under a confidentiality key K_C. The algorithm is based on AES-128 CTR or CBC mode. AES-128 is Advance Encryption System that produces 128-bit output from 128-bit input under the control of 128-bit key K_C [9].

To commercial wireless video applications, H.264/advanced video coding (AVC) standard has been widely chosen as video encoder because of its high compression rate and network friendly, so encryption algorithms should make full use of characteristics of the H.264 coded bit streams.

In addition, the compression and encryption process put on a heavy burden to mobile device, which has the limited power, processor resource and channel bandwidth. So encryption scheme should meet some special demands except for high security and low cost [8]:

1. Perceptual encryption: the encrypted video can still be decoded by any standard-compliant H.264 decoder, which is a basic feature of all perceptual encryption algorithms. Moreover, the original full quality video can be exactly recovered when the secret key is presented correctly.

2. Keep the compression ratio: the ultimate goal of video compression is to reduce the bitstream length to the minimum possible extent, so video encryption scheme cannot violate this fundamental goal.

3. Fast encryption speed: encryption scheme is desired to be used in the real-time applications, so the encryption should be easy to implement and avoid time-consuming.

4. Data operability: encryption scheme do not scramble the header and synchronization data for better data operability of encrypted bitstream.

Fig. 1 depicts the block diagram a selective video encryption at different levels of H.264/AVC. Five level encryptions are described in part III and are simulated and differentiated based on above demands.

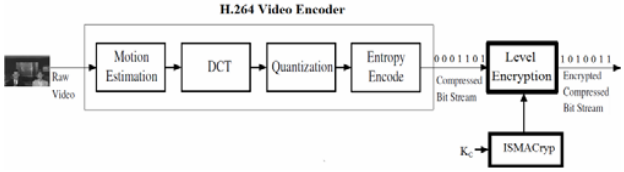


Fig. 1 Selective Video Encryption

II. H.264/AVC Overview

Jay M. Joshi is with the Electronics & Telecommunication Dept., S. V. M. nst. of Tech., Bharuch, 392001 India. (phone: +919426123589; e-mail: jaymjoshi@yahoo.com).

Upena D. Dalal is wElectronics Engineering Dept., Sardar Vallabhbhai National Institute of Technology, Surat, India. (e-mail: udd@eced.svnit.ac.in).

In this section, we introduce the necessary background knowledge about the H.264 standard that is closely related to video encryption work. In early 1998, the *Video Coding Experts Group* (VCEG) ITU-T SG16 Q.6 issued a call for proposals on a project called H.26L, with the target to double the coding efficiency (which means halving the bit rate necessary for a given level of fidelity) in comparison to any other existing video coding standards for a broad variety of applications. The first draft design for that new standard was adopted in October of 1999. In December of 2001, VCEG and the *Moving Picture Experts Group* (MPEG) ISO/IEC JTC 1/SC 29/WG 11 formed a *Joint Video Team* (JVT), with the charter to finalize the draft new video coding standard for formal approval submission as H.264/AVC [4] in March 2003. ITU-T Rec. H.264 | ISO/IEC 14496-10 version [5] refers to the integrated version 10 text after its amendment to define a new profile (the Constrained Baseline profile) intended primarily to enable implementation of decoders supporting only the common subset of capabilities supported in various previously specified profiles. In the ITU-T, the changes for versions 10 and 11 were approved on 16 March 2009.In this model, video encoding is applied sequentially, picture by picture, and for each picture, which is partitioned into one or more slices, encoding is applied to the slices independently. Each slice consists of a set of macroblocks, where, each is composed of one 16×16 luminance sample (Y) and two 8 × 8 chrominance samples (Cb and Cr) in the 4:2:0 or 4:2:2 chroma format [5]-[7].In this model of Baseline profile, slices are categorized into two types: I (Intra) slice and P (Predictive) slice. In H.264/AVC standard, the macroblocks in I slices are compressed without any motion prediction which is similar in earlier standards. In P slices, the macroblocks can be compressed with motion prediction. When using motion prediction, the macroblocks in P slices use one prediction from previous slice. For Intra macroblocks, spatial prediction is performed in H.264. To minimize the difference between the current block and previously encoded reference block, the 16×16 macroblocks are divided into three different block sizes: 16×16, 8×8, 4×4. Generally, blocks with smaller size, for example, 4×4, will offer better prediction efficiency than blocks with larger size, for example, 16×16, but will bring more overhead for mode decision bits. So it is suitable to use 16×16 blocks for luminance while 8×8 blocks for two chrominance. The reference block uses the block size 46×46 for luminance and 23×23 for chrominance. The prediction is based on SAD (Sum of Absolute Difference). The residual macroblock is encoded using CAVLC while other non visual contents like motion vectors are encoded using UVLC with exponential Golomb code. Slices in a picture are compressed by using the following coding tools:

- "Inter" spatial (block based) prediction
  - SAD used for motion estimation
  - Previous reference picture
  - I or P frame for reference
  - Fixed block size for motion estimation : 16×16 (Y) and 8×8 (Cb and Cr)
- 8×8 integer inverse transform (conceptually similar to the well-known DCT)
- Variable Scalar quantization (QP = 16)
- Zig-Zag (Frame) Coefficient scanning
- Lossless Entropy coding
  - Universal Variable Length Coding (UVLC) using Exp-Golomb codes for motion vectors
  - Context Adaptive VLC (CAVLC) for residual data
- Various color spaces supported (YCbCr of various types, YUV, RGB, etc.)
- 4:2:0, 4:2:2, and 4:4:4 color formats
- Flexible GOP (Group of Picture)
  - GOP 5- IPPPP
  - GOP 7-IPPPIPP

Table I summarize the result of simulation for two GOPs in terms of PSNR and Compression Ratio.

TABLE I

Name of Video	Resolution	FPS	GOP 7		GOP 5	
			PSNR (Y) (dB)	CR	PSNR (Y) (dB)	CR
Vipmen	120×160	30	33.60	21.27	33.89	16.71
Cat_video	120×160	30	31.12	32.70	31.70	21.45
Viptraffic	120×160	15	26.69	26.83	26.86	20.44
Vipcolorsegment	120×160	15	23.92	16.59	24.24	15.16
News	144×176	25	29.68	44.66	30.15	20.30
Vipbarcode	240×320	30	33.59	23.91	35.26	21.16
Vipfly	240×320	15	36.90	63.07	37.61	37.74
Vipmosaicking	240×320	15	35.36	28.46	36.93	24.43
Vipsnowydays	240×320	8	23.40	19.66	24.13	17.23
Vipunmarkedroad	240×320	1	29.24	20.54	29.78	18.36

SIMULATION RESULT OF H.264 CODING

We can encrypt different parts of NAL unit of H.264 to gain different security levels.

III. SELECTIVE ENCRYPTION USING ISMACRYP

Five levels of encryptions provided by this method are described as follows

- Level-5 : Encryption of all bits of NAL units (Naïve algorithm)
- Level-4 : Encryption of residual data (I and P frames)
- Level-3 : Encryption in nonzero and nonT1 levels of Quantized value in CAVLC
- Level-2 : Providing encryption of Level-3 in I frame only
- Level-1 : Encryption in MV (Motion Vector) only
- Level-0 : No encryption (Only coding)

A. Level-5 Encryption

Complete encryption algorithm provided by Level-5 which XOR video data as ordinary binary bit streams with secret keystream. It can use the CTR mode or CBC mode of block cipher algorithm AES-128 to enhance its security. Encryption the entire video data using standard encryption algorithms is referred as Naïve approach. The Naïve Algorithm encrypts the

entire video stream by treating it like a text stream. It's very safe due to all data encryption but it needs huge computational cost. Level-5 encryption encrypts data after compression and encoding, it doesn't change the compression ratio. But it doesn't have data operability because of header data encryption. The encrypted video by Naïve algorithm cannot be decoded by any standard-compliant H.264 decoder. Thus level-5 encryption is not perceptual encryption. Fig. 7 specify the encryption time of Level-5 in terms of Frame Rate Reduction ratio (% FRR) defined as

$$FRR = \frac{(\text{Frame rate without encryption} - \text{with encr}}{\text{Frame rate without encryption}} \quad (1)$$

Simulation results show that Encryption time is higher for Level-5 for GOP5 and GOP7. At this time, most of researches are searching selective video data encryption, which can reduce computational cost as it just encrypts only a part of video data. However the Selective Algorithm is not as secure as the Naïve Algorithm.

B. Level-4 Encryption

Level-4 based on the frame structure of H.264, encrypting the residual data of I frame and P frame only not the other parameter as shown in Fig. 2. The other parameter like image header, segment header, macroblock header data, Motion Vector etc. are remaining as it is. However huge amount of data are encrypted which increased computational cost and encryption time. Fig. 7 specifies that Level-4 does not provide much more improvement in terms of FRR compare to Level-5. As header data are not encrypted, it has the data operability. The Level-4 provides the synchronization but video cannot be played without decryption. It does not have perceptual encryption.

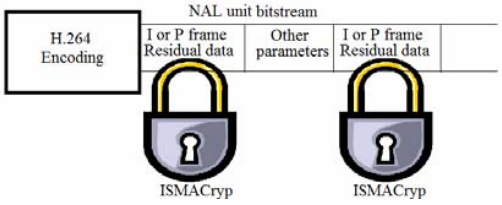


Fig. 2 Level-4 video encryption in I and P frame only

C. Level-3 Encryption

Level-3 encryption use the characteristic of entropy encoding –CAVLC [10]. It scans the coefficients in reverse order (from high frequency NZs to low frequency NZs) as shown in Fig. 3. CAVLC is designed to better exploit the characteristics of NZs, it works in several steps as shown in Fig. 4. Encoding of total NZs and number of trailing ones (T1's) is done by a single syntax element named *coeff_token*. It is followed by coding of signs of T1's. Remaining NZs are then coded using seven VLC tables. Lastly, total number of zeros and then runs of zeros are coded. To keep the bit stream compliant, which is a required feature for some direct operations (displaying, time seeking, cutting, etc.), we cannot encrypt *coeff_token*, total number of zeros and runs of zeros. The suitable syntax element which can be encrypted is the remaining NZs as shown in Fig. 4.

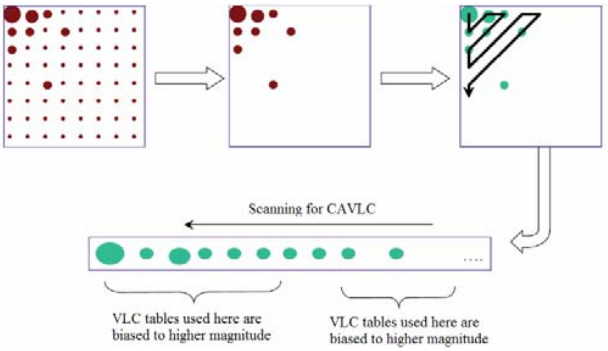


Fig. 3 Scanning order of NZs in CAVLC

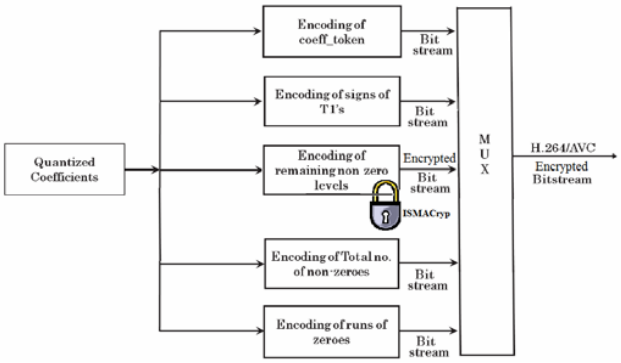


Fig. 4 Level-3 Encryption in CAVLC

It is perceptual encryption as shown by encrypted frame in Fig. 8(b) and has data operability capacity. Security is lower than Level-4 and Level-5, but it reduces computational costs and encryption time as shown in Fig. 7.

D. Level-2 Encryption

Level-2 encryption encrypts only I frames, because P frame is depends on I frame. Without decrypting I frame video decoder cannot decode the P frames. Level-2 encryption uses the Level-3 encryption only for I frame as shown in Fig. 5.

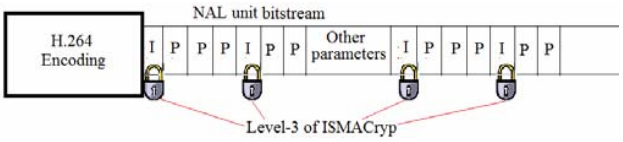


Fig.5 Level-2 encryption in I frame only

It is perceptual encryption as shown by encrypted frame in Fig. 8(c) and has data operability capacity. Security is lower than Level-3, Level-4 and Level-5, because it P-frames are unencrypted. But it further reduces computational costs and

encryption time as shown in Fig. 7.

E. Level-1 Encryption

Level-1 encryption scrambles only the Motion Vector using ISMACryp as shown in Fig 6. The 16×16 macroblock of luminance or 8×8 macroblock of chrominance contains only two values of MV(x,y) which is encoded using Expo- Golomb code. That contains very few bits in whole NAL unit, which provide lowest computational cost and encryption time as shown in Fig. 7. But the MV bits available only for P frames, thus I frames are decoded successfully without decryption. Thus Level-1 provides lowest level of security. The encrypted frames are shown in Fig. 8(d). In GOP 5 four frames are encrypted and first frame is unencrypted. In GOP 7 five frames are encrypted and two are not. However, the overall encrypted video become annoying for unauthorized customers. Thus, this type of encryption is well suitable for handheld device in Live TV application. The summary of all the levels is shown in table II.

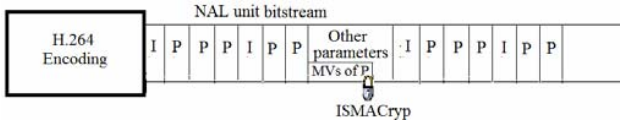
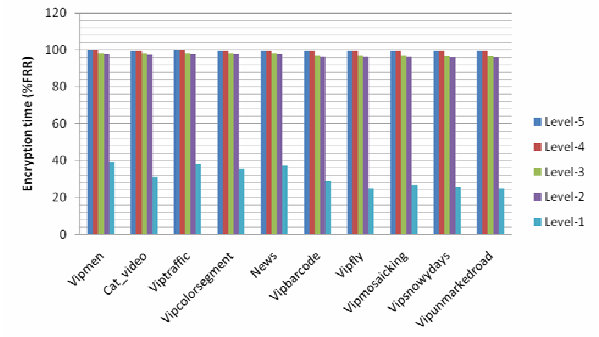
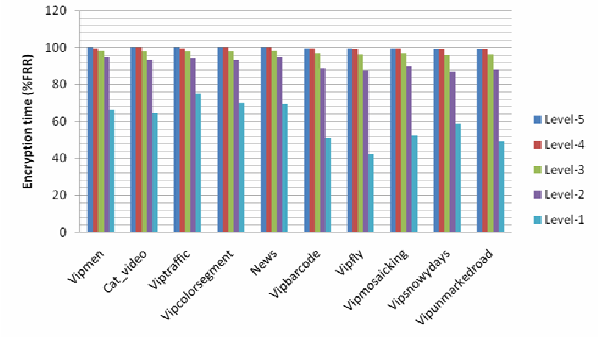


Fig. 6 Level-1 video encryption in MV only



(a)



(b)

Fig. 7 Encryption Time in terms of Frame Reduction Ratio for GOP (a) 5 and (b) 7

TABLE II
SUMMARY OF FIVE LEVELS ENCRYPTION

Encryption levels	Perceptual encryption	Change in compression ratio	Encryption speed	Data operability
Level-5	no	no	slow	no
Level-4	no	no	slow	yes
Level-3	yes	no	medium	yes
Level-2	yes	no	medium	yes
Level-1	yes	no	fast	yes

IV. CONCLUSION

This paper analyzes the performance of existing H.246-based video encryption algorithms and selective video encryption services' requirements, presents H.264-based multiple security levels scheme. In this paper, a perceptual and fast video encryption scheme Level-1 has been proposed specially for handheld device, which is based on exploiting the special feature of in H.264. Experimental results show that the proposed Level-1 scheme can achieve low security, fast encryption and low complexity cost without compromise on the compression ratio and transmission bandwidth. It is suitable for security multimedia services for mobile device and real time wireless application based on H.264 like Live TV, Video Chatting Video Conferencing etc. These schemes suitable for real time video transmission its hardware-implementation will be further studied in the future.

REFERENCES

[1] Stefan Doehla, Systems Engineer, Fraunhofer IIS , "DVB-H handheld video content protection with ISMA Encryption", pp 1-10, July 2007.
[2] ISMA Encryption and Authentication, Version 1.1; AREA / Task Force: DRM, September 2006.
[3] ISMA Encryption and Authentication Version 2.0; External Proposed Specification ; AREA / Task Force: DRM, November 2007
[4] "Draft ITU-T recommendation and final draft international standard of joint video specification (ITU-T Rec. H.264/ISO/IEC 14 496-10 AVC," in Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, JVTG050, 2003.
[5] ITU-T, "Advanced video coding for generic audiovisual services," ITU-T Rec. H.264; v11: March. 2009.
[6] Atul Puri, Xuemin Chen, Ajay Luthra, "Video coding using the H.264/MPEG-4 AVC compression standard", Signal Processing: Image Communication Vol. 19, No. 99, pp. 793-849, 2004.
[7] Iain E. G. Richardson: "H.264 and MPEG-4 Video Compression: Video Coding for Next-generation Multimedia", ISBN: 978-0-470-51692-8, John Wiley and Sons, Dec, 2003.
[8] Wang Li-feng, Wang Wen-dong, MA Jian, XIAO Chen and WANG Kong-qiao, "Perceptual video encryption scheme for mobile application based on H.264", The Journal of China Universities of Posts and Telecommunications: Science Direct, 15(Suppl.): pp 73–78, September 2008.
[9] Jay M. Joshi, Kiran R. Parmar and Upena D. Dalal, "Design and Implementation of KASUMI Algorithm in ISMACryp Encryption for Video Content Protection in DVB-H Application", IEEE International Conference on Control, Robotics and Cybernetics (ICCRC 2011), vol 1, ISSN: 978-1-4244-9709-6, pp 18-21, March 2011.
[10] Z. Shahid, M. Chaumont and W. Puech, "Fast Protection of H.264/AVC by Selective Encryption", WSPC – Proceedings : Singaporean-French IPAL Symposium, SinFra 2009, Fusionopolis, September 2009.
[11] Yan Li and Main Cai, "H.264-Based Multiple Security Levels Net Video Encryption Scheme", IEEE International Conference on Electronic Computer Technology: IEEE Computer Society, pp 8-11, 2009.
M. Abomhara, Omar Zakaria and Othman O. Khalifa, "An Overview of Video Encryption Techniques", IACSIT International Journal of Computer Theory and Engineering, Vol. 2, No. 1, pp 103-110, February, 2010.

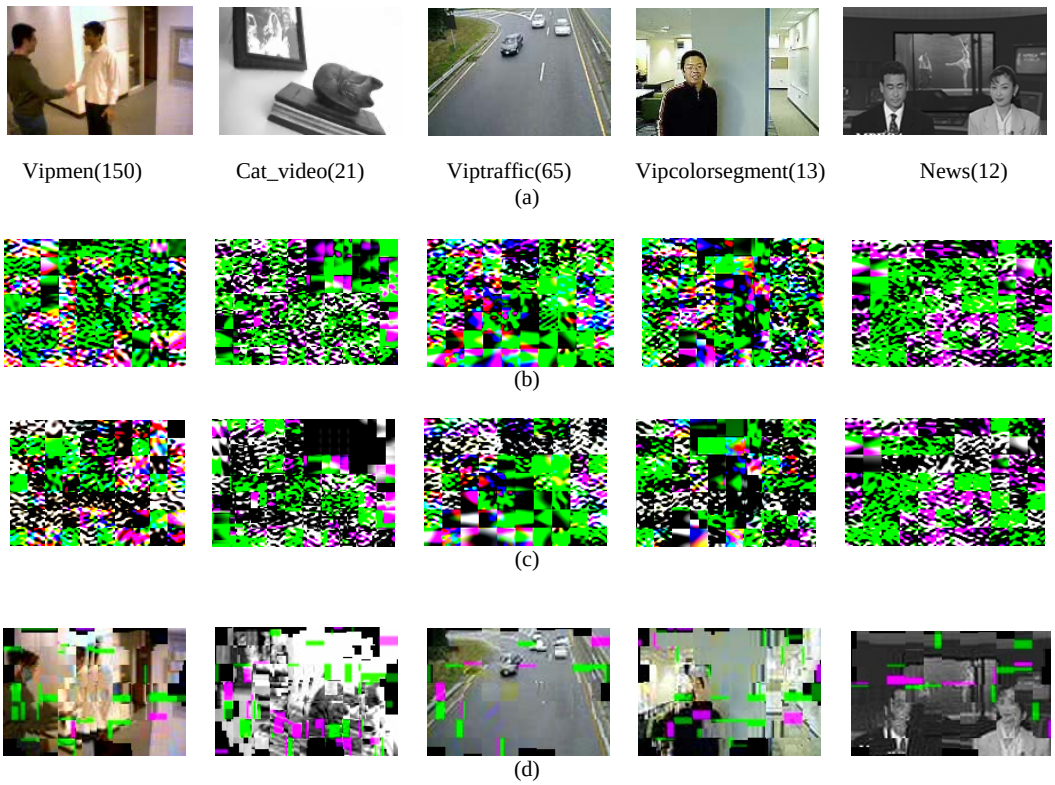


Fig. 8 (a) Original videos with their frame numbers. (b) Level-3 (c) Level-2 and (d) Level-1 encrypted frames.

Jay M. Joshi received his Bachelor's degree in Electronic & Communication Engineering from the Govt. Engg. College, Modasa in 2001. He obtained his Master degree in Communication System Engineering from L. D. College of Engg., Ahmedabad in 2003. He is pursuing Ph. D from S.V.N.I.T, Surat from July 2010. He is currently Associate Professor in department of Electronics & Telecommunication Engineering, S.V.M. Inst. Of Tech., Bharuch from last 7 years. He has total 8½ year teaching experience. His area of interest is Wireless Communication Systems, Digital Signal Processing, Digital image / video processing, coding and Compression, Wireless security and encryption. He published several papers in international journals and Conferences. He wrote several books on communication, Electronics and Signal processing areas. He is the active member of IEEE, IETE and ISTE professional bodies.

Upena D. Dalal received her Bachelor's degree in Electronic Engineering from the S.V.N.I.T, Surat in 1991. She obtained her Master degree in Electronics & Communication System from D.D.I.T, Nadiad in 2005. She obtained her Ph. D from S.V.N.I.T, Surat in 2009. She is currently Associate Professor in department of Electronics Engineering, S.V.N.I.T., Surat. She has 18 years teaching experience. Her area of research interest is Wireless Communication Systems, Video processing, coding and Compression, Wireless security and encryption. She published several papers in international journals and Conferences. She wrote several books on wireless communication areas. She is the active member of IEEE, IETE, Society of Women in Engineering, IJERIA, IE and ISTE professional body. She is also active member of social bodies like Surat Bhagini Samaj, Indian Menopause Society, Surat Mahila Club and Rashtriya Kala Kendra.