Improved Reversible Data Hiding Using Histogram Shifting Method

Arun K Mohan, Saranya M R
Department of Electronics Engineering,
School of Engineering & Technology,
Pondicherry University, Puducherry, India.
arunkannat@gmail.com, saranyamr.pu@gmail.com

K Anusudha
Department of Electronics Engineering,
School of Engineering & Technology,
Pondicherry University, Puducherry, India.
anusudhak@yahoo.co.in

***Abstract*— A reversible data hiding (RDH) algorithm with improved security, which can reacquire the cover in separable manner from the marked stego-image is presented in this paper. In the content owner side cover image is encrypted by deploying user-defined security key derived-chaotic based transposition algorithm. Then the data hider conceals secret data into the encrypted image by perturbing its histogram, by utilizing another user defined data hiding key. At the receiver side, the recuperation of the cover can be implemented directly or indirectly which depends on shared key. Lower bound of Peak Signal to Noise Ratio (PSNR) for direct recuperation method is set to 48.13dB. This technique has improved security & achieved higher data hiding capacity than the existing methods.**

Keywords*— *Reversible data hiding, Image encryption, PSNR, Chaotic sequence, Logistic Map Function, Histogram shifting, Key generation.

I. INTRODUCTION

In the last few decades we have witnessed the rapid development of communication technologies, global spread of the internet and the digital information revolution, which accelerated the use and transmission of multimedia information, broadened the scope of good and bad. Processing and transmission of multimedia information through highly insecure communication channels or networks causes predominant violation of internet privacy policies and becomes the attacks to the personal privacy. Since visual perception is more effective than the textual information, application of digital images are proliferated, now its security is imperative. To provide security attributes to image contents, we need to protect plain text from intruders. Image should be fortified from different type of attacks by introducing some safety mechanisms. In cryptography we usually employ encryption schemes to prevent any data from unauthorized access. But the normal data encryption schemes are not suitable for images.

In the very beginning of twenty first century researchers are attracted towards a new scheme called reversible data hiding (RDH) in encrypted images, RDH is a technique to embed a secret message into some distortion unacceptable cover media, like E-commerce Images, military, medical images, financial transactions, mobile phone applications, E-tendering and etc..[12], [13], in such a way that the cover image can be perfectly restored after unveiling the secret message. In order to disguise the of image information, the content owner encrypt the image before transmitting it to the receiver. RDH is gaining more significance among existing, since cover can be recovered without distortion after the embedded data is extracted while preserving the confidentiality of image content.

In literature, a plenty of excellent image encryption techniques are available by considering different aspects of image security. Image Encryption algorithms falls under two general categories: substitution based algorithms and transposition based algorithms. Some developers prefer both to enhance security. Substitution based encryption alters the grey level of the pixels to make the content concealed. A transposition based image encryption will not alter the pixel grey level, instead it shuffles the pixels in random according to some criteria. Permutation based encryption algorithms are covered in [1].

Researchers in steganographic domain have been proposing variety of schemes since the inception of RDH to perform data hiding techniques in encrypted image, [2] has adopted difference expansion technique. In this method, one bit can be embedded into two consecutive pixels. Consequently, the maximum embedding capacity is 0.5 bpp (bits per pixel). Later this method was generalized by Alatter with improved embedding capacity of (n-1)/n bpp. Different scheme of reversible data hiding, called reserving room before encryption is discussed in [3]-[4]. Other domain of RDH is histogram based method. [5]- [7] covers different methods under this domain. A new technique for RDH is implemented in [8] by estimating the errors. Methods in [3], [4], [5], [9] deals with RDH in separable Manner.

Data hiding is usually performed by assistant or channel administrator. In RDH content owner first disguise the image by employing an efficient encryption scheme. In such cases, when the owner needs to keep the secrecy of the image, by first encrypt the image using an encryption key. Then channel administrator, without any knowledge about the original image content, has to hide data into the encrypted image using a data hiding key. It is indispensable for receiver to maintain the separable reverse processing. Separable means, if the receiver is having only one key, then the corresponding reverse operation associated with that key can only be done, but not coupled or any fragment coupled operation. But irony is that if both keys are available to the receiver, then data extraction & image decryption are possible under any circumstances and in any succession. Proposed system possess all the features and

978-1-4799-1823-2/15/$31.00 ©2015 IEEE

requirements, implemented with the help of sensitive key derivation mechanism and secret message concealing algorithms based on the histogram amendment.

II. Proposed Method

We propose an efficient separable reversible data hiding in encrypted image with enhancement in image security by deriving initial seed parameters for the calculation chaotic sequence from the user specified key. Over all block sketch of proposed scheme is given in Fig.1.

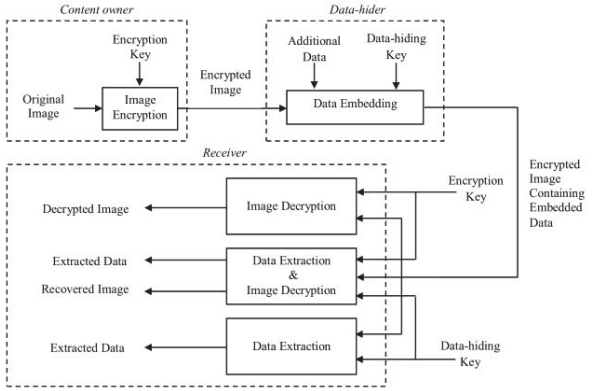

Fig.I. Over all block diagram of proposed scheme[5].

A. Image Encryption

Assume we are considering a grayscale image of size m × n. For a grayscale image, the pixel values are in the range [0, 255].

For image confusion chaotic based permutation method deployed in [1] is adopted. Logic of security enhancement is introduced by exploiting the strength of key derivation mechanism for initial seed reckoning of logistic map function with the help of 120 bit user defined key called encryption key. A sequence of pseudorandom numbers are generated by using the logistic function,

$$X_{P+1} = \mu \times X_P \times (1 - X_P) \quad (1)$$

Where X_P indicates the value of chaotic function at P^{th} iteration and the value of μ lies in the range [0, 4]. Usually initial seed of the logistic function lies between [0,1] but instead of directly specifying here algorithm prompts the user to enter the 120 bit encryption key and calculate the initial seed from the key using (2)-(5).

$$\text{Key} = \{k_1, k_2, k_3 \ldots k_{30}\} \text{ (In hexadecimal)} \quad (2)$$

Where k_i are the alphanumeric characters in the range (0-9) & (A-F). Each group of alphanumeric character indicates the section of the secret key, or this can be represented in ASCII code by giving notations from K_1- K_{15}. For the calculation of X_{01} consider $K_1K_8K_{15}$ and convert into binary (3),

$$X_{01} = (K_{11}\times 2^0 + \ldots + K_{18}\times 2^7 + K_{81}\times 2^9 + \ldots K_{88}\times 2^{15} + K_{15}\times 2^{16} + \ldots K_{158}\times 2^{23})/2^{24} \quad (3)$$

$$X_{02} = \text{Sum (ki)}/128 \quad (4)$$

$$X_0 = \text{Mean}(X_{01}, X_{02}) \quad (5)$$

Generate N = mn pseudorandom numbers using (1) and store them in a vector C (Chaotic). Sort the chaotic vector C in ascending order. A location map LM (Location Map) is generated by noting the changes in position during sorting. Location map is such that, LM (1) is the original index of smallest value, LM (2) that of second smallest and so on. Organize image pixels grey levels into a picture vector P (Picture). Rearrange P in accordance with the location map LM. Reshape picture vector P into matrix sized m × n to obtain the encrypted image.

B. Data Hiding

Since permutation based transposition algorithms for image confusion does not perturb the histogram of original image, histogram modification based method in [5] can be used to hide data into the encrypted image. It is suggested to have a clear idea about image histogram and image organization [10] is the prerequisite for not being perplexed. The method follows:

1. Pseudo randomly jumble the encrypted image pixels using data hiding key. Follow the same method as explained in the section II.A
2. First H (Header) pixels of binately confused cover are used to conceal parameters of data hiding. The parameters considered are MAX-MIN pixel values & number of MAX-MIN pairs. Let there are S MAX-MIN pairs. Then, the number of pixels needed to be reserved for header information is calculated using,

   $$H = (2 \times S + 1) \times MF \quad (4)$$

   MF is the multiplication factor for header which points towards representation index. Here grey level lies in the range [0,255] hence the multiplication factor to be chosen is eight.
3. Generate histogram of inert N-H pixels (here onwards inert pixels) which are not involved in the process of header generation. Find the maximum (MAX) and minimum (MIN) points. Here onwards a MAX point is the grayscale value having the maximum number of pixels in the image and a MIN point is the grayscale value having minimum number of pixels in the image.
4. Data hiding with a single pair only explained in this paper, one can reconfigure it according to the application and relevance. As per the observations and literatures, in most of the cases, MAX < MIN. In this method, bit stream of the secret message are embedded into binately confused cover pixels with grayscale value MAX.
5. Inert pixels used for generating the histogram are scanned in the sequential order. The grayscale values in the range [MAX+1, MIN-1] are provided with a unity accretion, by vacating the MAX+1. This indirect process of muddling the statistical parameter of the binately confused cover is called histogram shifting.
6. Let there be p pixels corresponding to the grayscale value MAX. These p pixels are used to hide the data, here onwards aliased by useful pixels. Secret message to be embedded is

converted to binary bit stream and stored in a vector called buffer. Binately confused cover is scanned for inert pixels in sequential order. Whenever a useful pixel is encountered, check the buffer for bit to be embedded. If the corresponding buffer bit is "1", then corresponding useful pixel grey level is supplied with unity accretion. If buffer bit is "0", useful pixel grey level kept unaltered.

7. Parameters of data hiding, number of MAX- MIN pairs and MAX-MIN pixel values, are hidden into binately confused cover by deploying LSB substitution on header pixels. The original H LSBs are also hidden as overhead information along with pure data.
8. Perform inverse permutation on the binately confused cover to get confused cover.

The eight steps described above will conceal the secret message into confused cover. It is observed that data hiding capacity of the method is (P-O-H) bits, where O is the overhead information due to MIN Positions. Improved data hiding capacity can be achieved by using multiple MAX and MIN pairs as explained in [11].

C. Data extraction & Image recovery

The two generalized possibilities in the real time practice, provided both keys are available to the receiver can easily be cracked out with the help of following three possibilities. The receiver is supplied with,

- Data Hiding Key alone
- Encryption Key alone
- Both Data hiding and Encryption Keys

1) Data Hiding Key alone

When receiver is supplied with data hiding key alone and confused cover containing concealed secret message, then only the extraction of concealed information alone is possible, but not the decryption of confused cover or any small fragment or region of confused cover.

For the unveiling of concealed secret message receiver should follow the concealing algorithm in the reverse order but in succession. First pseudo randomly permute the image to produce the binately confused cover on which we have implemented concealing algorithm. Follow the procedures in section II.B. Now move on to the parameter extraction from the header in the binately confused cover, extract the LSB of first eight pixels which signifies the number of MAX-MIN pairs S, then extract LSB of next 2×H× 8 header pixels which will give an idea about the MAX-MIN pair values. Receiver then scans the binately confused cover pixels in sequential order, whenever MAX + 1 is encountered, a bit "1" is extracted. If a pixel value MAX is encountered, a bit "0" is extracted. Bits extracted are concatenated with the help of buffer to interpret the concealed secret message. Likewise the complete concealed secret data as well as overhead information are extracted with full accuracy.

In some of the existing RDH techniques for restoring of original pixel values after extraction of the embedded secret message is improper. But in this we propose a scheme to recover the original pixel values after data extraction, complete binately confused cover after extraction is scanned again sequentially. LSBs of header pixels are replaced by the first H bits of extracted overhead information. For restoration of inert pixels we propose a unity curtailment to the grey levels in the range [MAX+1, MIN]. Replace these pixels with MIN value. After extraction of embedded secret message from the binately confused cover, perform inverse confusion using permutation to obtain confused cover.

2) Encryption Key alone

When receiver is supplied with encryption key alone and confused cover containing concealed secret message, then only the decryption of the cover alone is possible, but not the extraction of concealed secret message or any fragment of concealed secret message.

For this, receiver has to generate the same chaotic sequence using logistic map function as in encryption phase using equation (1) with the help of the sensitive key derivation mechanism to determine the initial seed. Location map which is used for confusing is generated by sorting the key derived chaotic sequence in ascending order. Using the location map, pixels of the confused cover are rejigged to obtain inverse confused cover. Process can bring back all the pixels to their original position. But distortion is introduced in the inverse confused cover by unity incongruity from primordial cover, in grayscale value for those pixels used for embedding the binary value "1" of the secret data buffer. It can be prove that lower bound of Peak Signal to Noise Ratio(PSNR) is fixed to 48 dB.

It is observed that the grayscale value of pixels between MAX and MIN will either be incremented or kept same during data hiding. In the worst case, the value of every pixel differs by a value of 1 from their original value. Thus, the Mean Square Error (MSE) of worst case is 1 and the lower bound of PSNR of the decrypted image containing hidden data is given by,

$$\begin{aligned} PSNR &= 10 \log_{10} ((255\times255)/MSE) \\ &= 10\log_{10} (255^2) \\ &= 48.13 \text{ dB} \end{aligned}$$

This resultant lower bound of PSNR is much higher than that of reversible data hiding in encrypted image techniques, [12] and [2], reported in the literature.

3) Both Data hiding and Encryption key

When the receiver has both keys, then both the operation explained in section C.1 & C.2 can be done in any order. Hence the condition of separability also satisfied.

III. Simulation Results

We coded and tested the proposed separable RDH with enhanced image security in Matlab R2012b with 4 commonly used test images namely Cameraman, Lena, Rice and Tulips, each of size 256×256. In our implementation we used two peak points of histogram for concealing secret message. When the receiver is supplied with encryption key alone, then image can be decrypted and the PSNR of all directly decrypted images were observed to be above 48.13dB and verifies the theoretical

result. When receiver is supplied with data hiding key alone, then concealed secret message can be extracted without any error. When the receiver is supplied with both encryption key and data hiding key, then both the operations can be done in separable manner. The recovered image is exactly same as the original image as in [3]. Fig.II shows the images during different phases of the algorithm.

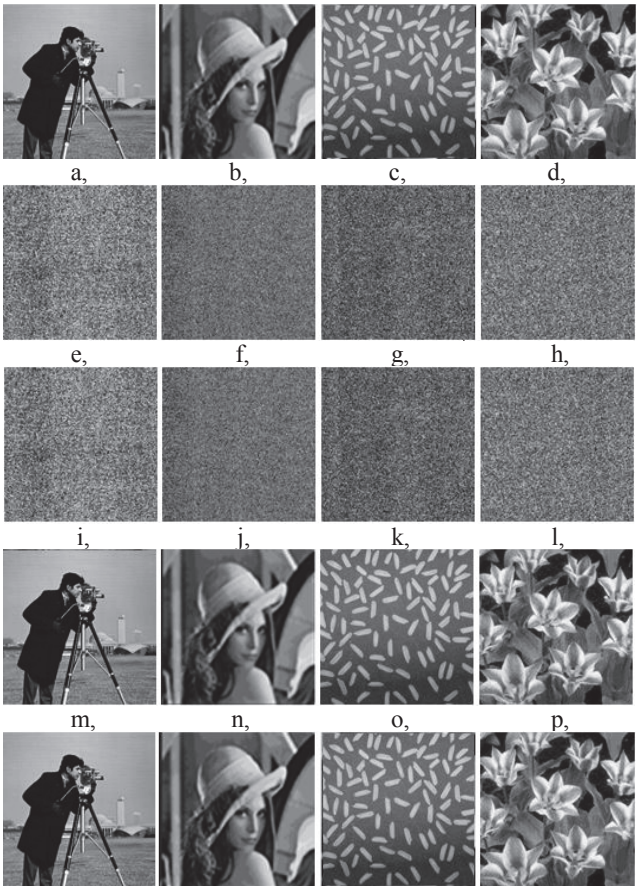

Fig.II. Test Image during different stages of proposed algorithm

a-d) Input Images(Cameraman, Lena, Rice & Tulips)
e-h) Encrypted Images(Cameraman, Lena, Rice & Tulips)
i-l) Encrypted Images data hidden(Cameraman, Lena, Rice & Tulips)
m-p) Directly Decrypted Image(Cameraman, Lena, Rice & Tulips)
q-t) Decrypted after extration of data(Cameraman, Lena, Rice & Tulips)

Table I. shows the data hiding rate in bits per pixel (bpp) with and without restriction, *ie,* embedding rate is calculated using application and number of pixels in first two peeks, PSNR of directly decrypted images. Table II gives the small comparison between existing and proposed system in terms of PSNR for directly decrypted images.

TABLE.I, Experimental Results for some commonly used images with & without restriction, embedding rate is indicated in brackets.

Input Image	PSNR of recovered Image (db)	PSNR of recovered Image (db)
Cameraman	53.1814(0.0241)	70.3961(0.005)
Lena	52.9996(0.0173)	68.6688(0.005)
Tulips	50.6855(0.0204)	70.4268(0.005)
Rice	48.4735(0.0285)	68.4069(0.005)

Table II. Comparison of PSNR values of Existing and Proposed Systems

Input Image	Existing System	Proposed System
Cameraman	49.6220	70.3961
Lena	53.7269	68.6688
Rice	51.2141	70.4268
Tulips	48.6855	68.4069

IV. Security Analysis

Under this we discuss some parameters which are used to ensure the required security to the marked stego image.

A. Correlation coefficient analysis

Correlation coefficient is a measure for statistical analysis which gives the correlation between adjacent pixels in the image. Less correlation coefficient indicates stronger ability to resist statistical attack. Here a number of pairs of adjacent pixels are randomly selected from the plain image and the encrypted image to calculate the horizontal, vertical and diagonal correlation coefficient. Table III shows the results of correlation coefficient analysis, and it is clear that the correlation between adjacent pixels in the image has been significantly reduced after encryption.

Table III. Correlation Coefficient Analysis

		Cameraman	Lena	Rice	Tulips
Original Image	Vertical	0.9943	0.9798	0.9231	0.9726
	Horizontal	0.9903	0.9893	0.8660	0.9568
	Diagonal	0.9854	0.9697	0.8543	0.9343
Marked Stego Image	Vertical	0.0033	0.0090	0.0023	0.0003
	Horizontal	-0.0017	0.0006	0.0004	0.0032
	Diagonal	0.0042	0.0024	0.0018	0.0003

B. Key space analysis & Key Sensitivity

It is the total number of different keys that can be used in an encryption system. Key space should be sufficiently large to make brute force attack infeasible. In the proposed system we deployed 120 bit key for the encryption, which is used as the set of 15 eight bit keys supplied to an algorithm for generating initial seed, the order of application of grouped key is also influencing the security. So the possibility of guessing the key is once only in 2^{120}. It may take few years to get success if the key space is large enough. Suppose if the intruder deploy a dedicated computer performs 1000MIPS, then the year of load of that computer can be calculated using.

$$\frac{2^{120}}{1000\times10^{6}\times60\times60\times24\times365} \cong 4.21495432453 \times 10^{19} \text{ Years}$$

This is very long period and practically infeasible.

By using the 120bit key, we tested and verified the single digit change in the key will provide the complete random image, *ie*, not revealing any of the information about the encryption. So the proposed system is completely sensitive to the key.

C. Brute force attack

In brute force attack, cryptanalyst tries all possible keys in a finite key space one by one in the sequential order and check the corresponding plaintext, if meaningful that key will be chosen for decrypting entire data . By literatures average of,

half of all possible keys must be tried to achieve success, For this attack a large computation is involved which will introduces high complexity. Hence this attack is not possible for the proposed system by exploiting the large key space. Table IV shows time involved for various key space [17], [18].

TABLE.IV. Average time required for exhaustive key search

Key size(bits)	Number of alternative Keys	Time required at 10^6 decryptions/μs
32	2^{32}	2.15 milliseconds
56	2^{56}	10 Hours
128	2^{128}	5.4×10^{22} Years
256	2^{256}	5.9×10^{30} Years

D. Differential attack

General requirement of an encryption scheme is that encrypted image and the original one should be greatly different. Difference can be measured using two parameters called, number of pixel change rate (NPCR) and the unified average changing intensity (UACI) [19]. During differential attack, a minute change to the original image is made by intruder using the introduced algorithm to encrypt the main image before and after changing and by comparing two encrypted images to realize the correlation between the original image and the encrypted one. Correlation is a measure that computes degree of similarity between two. Its coefficient is a useful in measuring encryption quality of any cryptosystem. A cryptosystem is said to be good only if encryption algorithm conceals all attributes of a plaintext image, and encrypted image is totally random which means plaintext image and original image are highly uncorrelated. The above said parameters are used to measure the immunity to the differential attacks. Higher value of NPCR & UACI will provide high immunity, while correlation coefficient signifies the measure of randomness. Table V shows these measures of the proposed system.

TABLE.V, Obtained NPCR, UACI, & Correlation coefficient

Input Image	NPCR	UACI	Correlation Coefficient
Cameraman	0.9961	0.3991	0.0124
Lena	0.9929	0.3318	0.1732
Tulips	0.9943	0.3681	0.3211
Rice	0.9959	0.3832	0.2876

V. CONCLUSION

We proposed a novel algorithm for enhanced image security with reversible data hiding in a separable manner. The algorithm is further divided into several modules like image encryption, data hiding, data extraction and image recovery phases. In image encryption phase, content owner encrypts the cover by deploying user defined key derived chaotic based transposition algorithm. In the second phase data hider without knowing the cover content, conceal data into the encrypted image by deploying an algorithm based on data hiding key derived chaotic sequence. The advantage of this method is that it can adopt configurable data hiding rate which suites this for E-Commerce, E-tendering and Medical Image applications. At the receiver, data extraction and image recovery are performed in a separable manner in accordance with the key supplied to them. The lower bound PSNR of this decrypted image is 48.13 dB, which is much higher than that of existing reversible data hiding techniques in encrypted image. The proposed method is analyzed for understanding the immunity towards different attacks and proven immune.

VI. REFERENCES

[1] Chinmaya Kumar Nayak, Anuja Kumar Acharya, Satyabrata Das, "Image encryption using an enhanced block based transformation algorithm", International Journal of Research and Reviews in Computer Science, vol. 2, No. 2, Apr. 2011.

[2] Jun Tian, "Reversible data embedding using a difference expansion", IEEE Trans. on Circuits and Systems for Video Technology, vol. 13, No. 8, Aug. 2003.

[3] K. Ma, W. Zhang, X. Zhao, N. Yu, F. Li, Reversible data hiding in encrypted images by reserving room before encryption, IEEE Trans. Information Forensic Security. 8(3) (2013) 553-562.

[4] W. Zhang, K. Ma, N. Yu, Reversibility improved data hiding in encrypted images, Signal Process. 94 (2014) 118–127.

[5] Arun K Mohan, Saranya M R, K Anusudha "Algorithm for Enhanced Image Security with Reversible Data Hiding", International conference on Contemporory Computing & Infomatics, Nov, 2014.

[6] Mohammad Ali Bani Younes, Aman Jantan, "An image encryption approach using a combination of permutation technique followed by encryption", International Journal of Computer Science and Network Security, vol.8 No.4, April 2008.

[7] Hiral Rathod, Mahendra Singh Sisodia, Sanjay Kumar Sharma, "Design and implementation of image encryption algorithm by using block based symmetric transformation algorithm", International Journal of Computer Technology and Electronics Engineering, vol. 1, Issue 3.

[8] Weiming Zhang, Kede Ma, Nenghai Yu,"Reversibility improved data hiding in encrypted images", Elsevier Signal Processing letter, vol. 94, (118-127), Jun 2014.

[9] Xiaotian Wu, Wei Sun, "High capacity RDH in Encrypted Images by prediction error", IEEE Signal Processing Letters, vol. 18, No. 4, Apr. 2011.

[10] Digital Image Processing By Rafael C. Gonzalez Edn 2, Chapter 2.

[11] Xinpeng Zhang, "Separable reversible data hiding in encrypted image" IEEE Trans. on Information Forensics and Security, vol.7, No. 2, Apr. 2012".

[12] B. Acharya, S. Patra, and G. Panda, "Image encryption by novel cryptosystem using matrix transformation," in Emerging Trends in Engineering and Technology, 2008. ICETET'08. First International Conference on. IEEE, 2008, pp. 77–81.

[13] G. Jakimoski and K. Subbalakshmi, "Cryptanalysis of some multimedia encryption schemes," Multimedia, IEEE Transactions on, vol. 10, no. 3, pp. 330–338, 2008.

[14] W. Zeng, H. Yu, and C. Lin, Multimedia security technologies for digital rights management. Academic Pr, 2006.

[15] J. Zhou, O. Au, X. Fan, and P. Wong, "Joint security and performance enhancement for secure arithmetic coding," in Image Processing, 2008,ICIP 2008. 15th IEEE International Conference on. IEEE, 2008, pp. 3120–3123.

[16] B. Schneier, Applied Cryptography. John Wiley & Sons, Inc., USA, 1996.

[17] W. Stallings, Cryptography and network security: principles and practice. Prentice Hall, 2010, vol. 998.

[18] C. Li, S. Li, D. Zhang, and G. Chen, "Cryptanalysis of a data security protection scheme for voip," in Vision, Image and Signal Processing, IEE Proceedings-, vol. 153, no. 1. IET, 2006, pp. 1–10.

[19] C. Shannon, "Communication theory of secrecy systems," Bell system technical journal, vol. 28, no. 4, pp. 656–715, 1949.