

AN ADAPTIVE PSEUDORANDOM STEGO-CRYPTO TECHNIQUE FOR DATA COMMUNICATION

Ajay Kumar Nain¹, Shailender Gupta², Bharat Bhushan³, Rashmi Chawla⁴

YMCA University of Science and Technology

¹aknain90@gmai.com, ²shailender81@gmail.com, ³bhrts@yahoo.com,

⁴rashmi.chawla@rediffmail.com

ABSTRACT

To provide security in data communication networks various forms of cryptographic and steganographic algorithms have been proposed. Though both of these techniques serve the same purpose i.e. to secure data but have some drawbacks. Using cryptography might raise some suspicion whereas in steganography the distortion in cover can attract an adversary. The combinations of the two as proposed by various researchers can be very fruitful in terms of brute force search time and robustness against attacks. A popular method of combination is using pseudorandom Least Significant Bit (LSB) substitution along with stream cipher or block cipher. The problem with pseudorandom LSB substitution technique is with the random pixel selection process. It inserts the cipher text on randomly chosen pixels using random interval method having certain period. Now if this period of random interval number generations is large, it increases time complexity. On the other hand if it is too low then the cipher text will not be evenly distributed in all the colour planes. This paper proposes an adaptive method on the basis of size of secret data for random pixel selection for data embedding. The results show that using such an approach not only increases brute force search time but also reduces the time complexity of the overall process.

KEYWORDS

Steganographic noise, adaptive colour image steganography, random pixel selection, histograms of colour image, Difference in histograms, MSE, PSNR

1. INTRODUCTION

Now-a-days the need for information security has increased with the advance in internet computing. For this purpose various cryptographic and steganographic techniques have been proposed in literature. Cryptography [2] is basically used to secure the confidential information by encrypting the secret message using encryption and decryption. On the other hand Steganography [1,5] is the art of hiding information within innocuous cover carriers in such a way that the hidden message is undetectable. Though the purpose of both the techniques is same but both the techniques have limitations. If we had encryption protection in place, an opponent might still be able to observe the pattern of the transmitting messages. The opponent could determine the location and identity of communicating hosts and could observe the frequency and length of messages being exchanged. This information might be useful in guessing the nature of the communication that was taking place. Similarly in steganography the changes in statistical properties [8, 11, 12] of the cover can be used by opponent to detect the presence of secret message.

To overcome the above problem researchers have proposed combination of the steganography with cryptography [14, 17]. It provides better security to cipher text. In this case if somehow

presence of message is detected by an opponent, it requires an additional effort to decrypt the message also. A popular method available in literature is combining pure LSB substitution technique along with stream cipher or block cipher. The LSB substitution steganography inserts the cipher text in the LSB of pixel value [9, 10]. This technique is popular due to presence of superfluous information in image which can be easily interpolated [3-4]. To increase the brute force search time the LSB technique is replaced by pseudo random LSB method. In this method random pixels [6, 7] are selected on the basis of seed value and LSB at corresponding pixels is replaced by cipher text. The problem with pseudorandom LSB substitution technique is with the random pixel selection method. This technique inserts the cipher text on randomly chosen pixels using random interval method having certain period. Now if this period of random interval number generations is high, it increases time complexity. On the other hand if the period is too small then the cipher text will not be evenly distributed in all the colour planes.

This paper proposes an adaptive method taking care of above problems. It combines stream cipher cryptographic technique with modified pseudorandom LSB substitution technique to provide evenly distribution of cipher text. In addition to it, the proposed method provides enhance security in terms of brute force search time. It also reduces time complexity value by avoiding the number of collisions.

The rest of the paper is organized as follows: Section 2 provides the literature survey and the problem identification. Section 3 provides the algorithm of the proposed technique. The simulations set up parameters, performance metrics taken are given in section 4. Section 5 gives the results and discussion followed by concluding remarks and references.

2. LITERATURE SURVEY

Some researchers who have previously worked in this direction as follows:

Shailender Gupta et.al. [14, 16] proposed a hybrid model that combines LSB technique with symmetric or asymmetric key cryptography. The results show using a hybrid approach that combines asymmetric key cryptography with steganography is not viable. On the other hand the hybrid approach has an advantage of increased brute force search time. The problems with this approach can be summarized as follows:

- The cipher text is not evenly distributed in all the colour planes.
- The brute force search time of this technique is quite high.
-

Recently Ajay Nain et.al. [17] proposed a technique that combines pseudorandom LSB with RC4 cryptography mechanism. Their approach not only improved the brute force search time of the overall system but an effort was also made to randomly distribute the cipher text in all the colour planes. The problem with this approach can be summarized as follows:

- The distribution of cipher text in all the colour planes was not proper
- The time complexity of the overall system increased due to increase in number of collisions.

This paper has been inspired from the above mentioned literature. To improve the above hybrid techniques the following aspects were kept in mind:

- To evenly distribute Cipher text in all the colour planes.
- To increase the brute force Search time of overall system.

- To reduce the value of time complexity
- To minimize the number of collisions during random pixel selection

The next section gives the proposed technique to achieve the above mentioned objectives.

3. PROPOSED TECHNIQUE

Before discussing the proposed algorithm we would like to give a brief description of the variables and functions (see Table 1) used in our proposal that will help the readers in better understanding of the proposal.

Table 1. Terminology used in the paper.

Entity	Description
C	Original Image
S	Stego-Image
P	Plain Text
D	Encrypted secret data
L(d)	Length of encrypted data
M * N	Image dimensions
Seed	seed used to generate random numbers
Random()	Function to generate uniformly distributed random numbers
Rc4_Encrypt()	Encrypt data using DES or RC4
Rc4_Decrypt()	Decrypt data
PPN	Pixel position number
Key	Key for cryptography
kn	Number of bits used for LSB substitution
a, b, c	Parameters to be set experimentally

The main step of the proposed technique is to identify the pixels of the cover image where the cipher text bits are to be stored. This process is same for both the sender and receiver. For this purpose a Pixel Position Number (PPN) array is calculated in which each member indicates a unique pixel in the cover image. We denote this process of generating PPN by function *Random()* whose inputs are seed, length of cipher text and parameters a, b and c. The process of its generation is explained below:

The first step is to calculate the number of LSB (kn) require for substitution depending upon the length of secret message L(d) and size of cover image(MxNx3).

$$k_n = \text{ceil} \left(\frac{a \times L(d)}{M * N * 3} \right)$$

Where *ceil()* is the greatest integer function

The next step is to identify the pixels of the steganographic cover (S) where the cipher text bits are to be embedded. For this purpose first of all random numbers are generated between 1 and max where max is calculated as

$$max = b * floor\left(\frac{(M \times N \times 3) * k_n}{L(d)}\right)$$

Where *floor()* is the smallest integer function

Parameter b decides the number of round to complete selection of pixels over the red, green and blue plane of cover image. The setting of parameter b is discussed in section 5. From this array (B) of random numbers PPN is calculated by adaptively addition of random numbers as given below:

Algorithm: Function Random ()

```

n=B(1); //n is a temporary variable
PPN(1)=n;
for i=2 to (L(d)/kn)
    if round >2
        mp=1;
    else mp=0;
    end if
    n=(n+B(i)+c*mp*B(i)) %(MxNx3)
    if value of n already exist in PPN //check for collision
        n=next value;
    end if
    PPN(i)=n;
end for
output: PPN
    
```

In this method we accelerate the width of random interval after completion of two rounds of pixel selection over entire range of cover image. The value of parameter c decides the change in interval. The adjustment of this parameter is discussed in section 5. To overcome the problem of collisions, we compare the current value of PPN from all the previous value of PPN and if the value is found prior, it means collision is detected. In that case we discard the current value and choose next value of n. Otherwise we store the current value in PPN array.

The rest of the process of the proposal is explained separately for sender and receiver.

3.1. Sender Side Procedure

The plain text (P) data to be transferred is read as a set of ASCII characters at the sender side. This text is encrypted using function Rc4_Encrypt () and key. Encrypted text (D) is represented in bit-stream form and is denoted by d. Then PPN is generated by calling Random() function as discussed above. Corresponding to each value of PPN the location of pixel in image is identified by calculating the height (m), width (n), and plane (p) of pixel respectively. Then kn LSBs of corresponding pixel is replaced by the encrypted data bits. In this way the encryption process is completed as explained in algorithm:

Algorithm: Embedding Process (sender side)	Algorithm: Extracting Process(receiver side)
Input: C, Key, seed, P S=C;	Input: S, Key, seed, L(d) $k_n = \text{ceil}(a * L(d) / (M * N * 3));$

<pre> D=Rc4_Encrypt(Key,P); d=bit-stream(D); kn=ceil(a*L(d)/(M*N*3)); PPN=Random(b, c, L(d), seed); for i=1:L(d)/kn t1=PPN(i); p=ceil(t/(M*N)); t2=t1-(p-1)*M*N; m=ceil(t2/N); n=t2-(m-1)*M; S(m,n,p) <= d(i,i+kn); End for </pre>	<pre> PPN=Random(b, c, L(d), seed); for i=1:L(d)/kn t1=PPN(i); p=ceil(t/(M*N)); t2=t1-(p-1)*M*N; m=ceil(t2/N); n=t2-(m-1)*M; d(i,i+kn) <= S(m, n, p); End for P=Rc4_Decrypt(key, D); </pre>
Output: S	Output: P

3.2. Receiver Side Procedure

For extraction of plain text receiver must know the seed and length of the message L(d) so that to identify the embedded pixels in stego- image and Key to decrypt the message can be identified. First of all PPN is generated in the same way as explained above. The kn LSBs of these pixels are taken as data and then this data is decrypted using Rc4_Decrypt algorithm to obtain plain text (P).

4. SIMULATION SETUP

4.1. Performance metrics

4.1.1. Mean Square Error (MSE)

It measures the statistical difference between the cover and stego-image. The distortion in the image can be measured using MSE which can be calculated using Equation 1.

$$MSE = \left[\frac{1}{M * N} \right]^2 \sum_{i=1}^M \sum_{j=1}^N (X_{ij} - X'_{ij})^2 \quad \dots\dots(1)$$

Where:

X_{ij} : The intensity value of the pixel in the cover image.

X'_{ij} : The intensity value of the pixel in the stego image.

M*N: Size of an Image.

4.1.2. Peak Signal to Noise Ratio (PSNR)

It is the measure of quality of the image by comparing the cover image with the stego-image. High PSNR indicates good perceptual quality of stego-image. It is calculated using Equation 2.

$$PSNR = 10 \log_{10} \frac{255^2}{MSE} \text{ db} \quad \dots\dots(2)$$

4.1.3. Histograms

Histogram is a measure of the number of occurrence of pixels with respect to particular pixel value [13]. During embedding pixel value changes, hence number of pixels having a particular pixel value gets changed. These changes can be used to detect steganography.

4.1.4. Steganographic Noise

It shows the pixels in the image on which data is embedded. Evenly distribution of noise all over the image indicates good quality of steganography technique.

4.1.5 Collisions

During identification of pixel positions if a pixel position number is repeated more than once then it is referred to as collision. These collisions adversely affect time complexity of the process.

4.1.6 Time Complexity

The total time taken by all the processing on receiver and sender side comprises of time complexity.

4.2 Setup parameters

The MATLAB Version 7.13.0.564 (R2011b) is used to implement and simulate the proposed technique. MATLAB is used because of large number of advanced inbuilt functions and image processing toolbox. The various simulation parameters are given in Table 2.

Table 1. Setup parameters.

Cover image	256*256*3 (colour)
Image type	Bmp
Secret text file size (kb)	0.5, 2 ,8,32
Simulation Tool	MATLAB 7.13.0.564
Pseudo Random Number Generator	'Uniform' with seed=33559
Cryptography	RC4 key= 'ajaynain'
Processor	Core-i3 @2.53 Ghz RAM 3GB

5. EXPERIMENTAL STUDY

5.1. Adjustment of parameters

We have used three parameters a, b and c in the proposed method. To accurately measure the performance of the proposed technique it is important to optimize these parameters. For optimization purpose we analyzed the time complexity, collisions, distribution of steganographic data and complexity of hidden message.

- The value parameter ‘a’ is inversely linked to embedding capacity. If value of this parameter is low then embedding capacity and distortion is high and time complexity is low. To find the value of ‘a’ experiments were performed and the best results were obtained for value equal to 2.
- The value of ‘b’ decides the number of rounds to allocate the random pixels for embedding the cipher text evenly throughout the image. If the value of ‘b’ is kept high then number of rounds would increase resulting in increased number of collisions and time complexity. On the other hand if value of ‘b’ is taken small then the complexity of hidden message is reduced because of adjacent bits of message comes more close on the cover image. Value of ‘b’ as 4 is considered as an optimum one.
- The value of ‘c’ scatters the remaining bits after two rounds of pixel selection over the full range of image. The value of ‘c’ is set to 6 in case of our experiment since at this value scattering of cipher text was perfect.

5.2. Results

5.2.1. Perceptual Quality

By inspection of Figure 1 we can say that proposed techniques does not introduce significance changes in the image quality.





File size =.5kb		
	Figure 1(a) Impact of Proposed Technique	Figure 1(b) Impact of Previous Technique
File size =2 kb		





	Figure 1(c) Impact of Proposed Technique	Figure 1(d) Impact of Previous Technique
File size =8 kb		
	Figure 1(e) Impact of Proposed Technique	Figure 1(f) Impact of Previous Technique
File size =32 kb		
	Figure 1(g) Impact of Proposed Technique	Figure 1(h) Impact of Previous Technique

Figure 1. Impact on Perceptual Quality

5.2.2. Histogram Representation

The differences in histograms of cover image and the image after applying previous and proposed technique are shown in Figure 2. The Figure 2(a) shows the histogram of the actual image while the Figure 2(b) and Figure 2(c) shows the impact on histogram after applying previous and proposed technique on image. Figure 2(d) to Figure 2(k) show the differences in the number of pixel values between original image and the image obtained after applying the proposed mechanism. The following inference can be drawn from the Figure 2:

- As the size of secret file increases the changes in histogram increases.
- In proposed technique the changes in histogram are almost same in the red, green and blue colour planes while in the previous technique the cipher text is only distributed in Red plane only.

Red Plane

Green Plane

Blue Plane

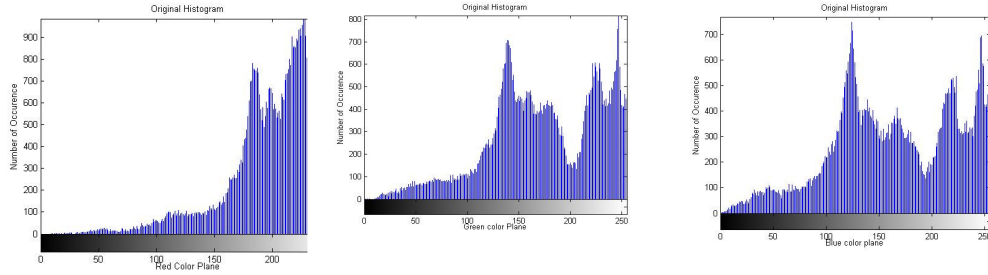


Figure 2(a). Histograms of original cover image

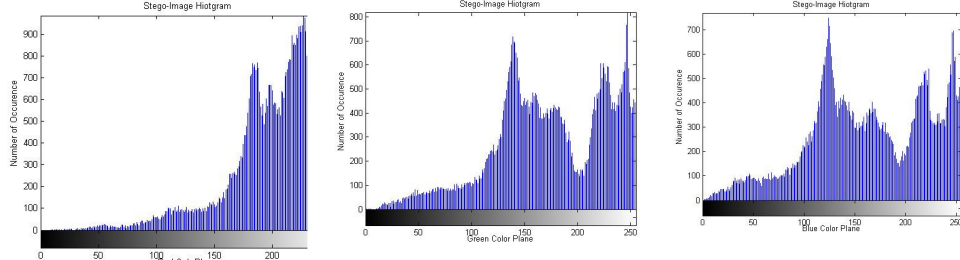


Figure 2(b) Histograms of stego-image of previous technique

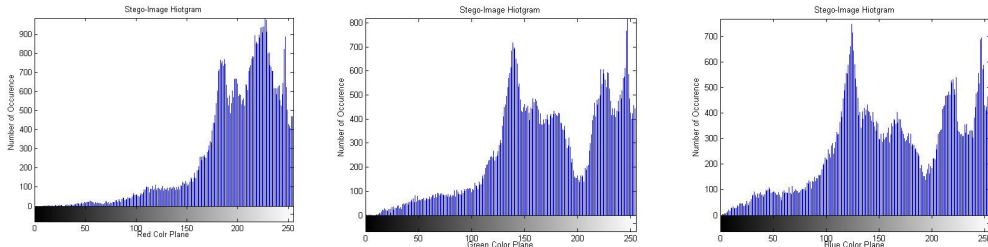


Figure 2(c) Histograms of stego-image of proposed technique

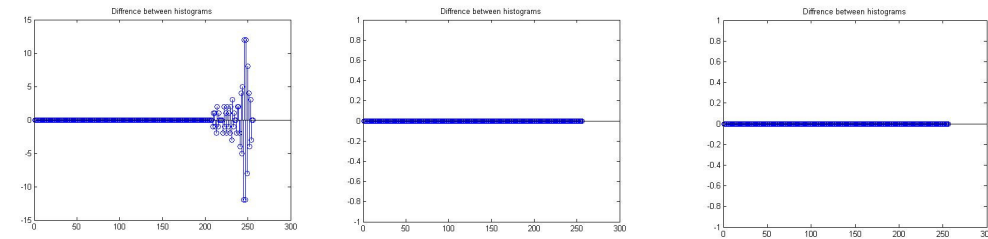


Figure 2(d) Difference in Histograms for previous technique with file size =0.5 kb

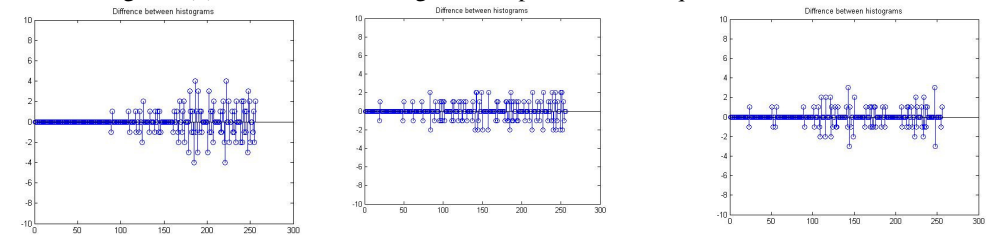


Figure 2(e) Difference in Histograms for proposed technique with file size = 0.5 kb

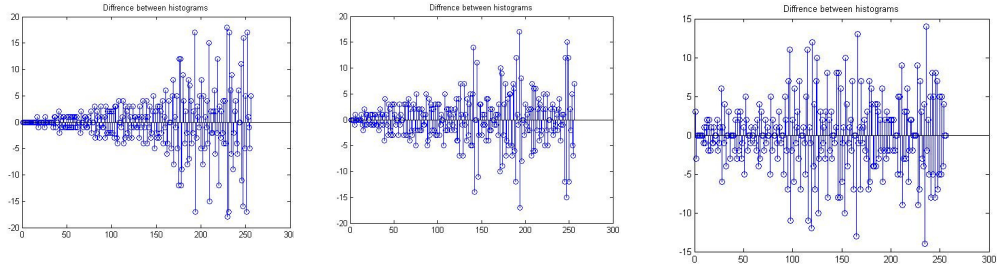


Figure 2(f) Difference in Histograms for previous technique with file size =2 kb

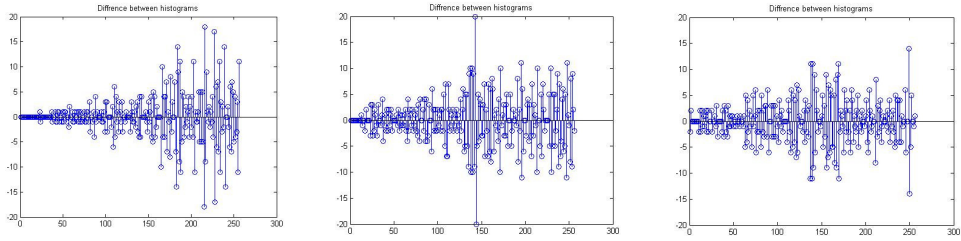


Figure 2(g) Difference in Histograms for proposed technique with file size=2 kb

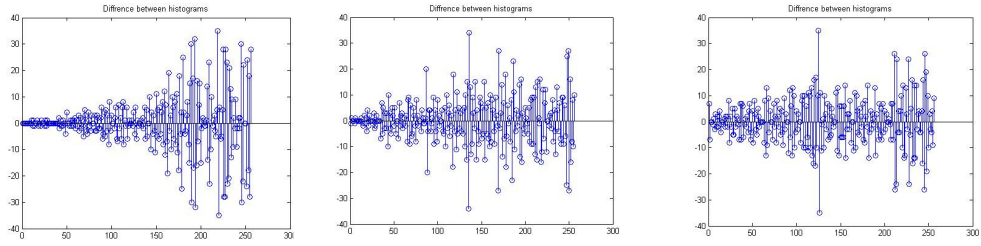


Figure 2(h) Difference in Histograms for previous technique with file size =8 kb

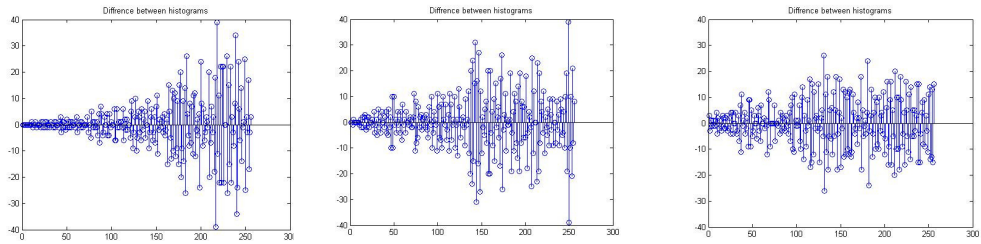


Figure 2(i) Difference in Histograms for proposed technique with file size =8 kb

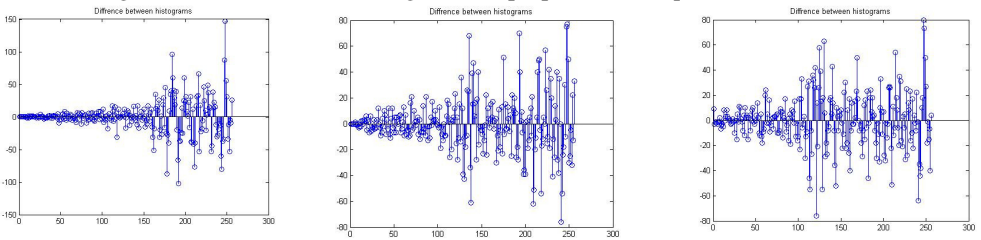


Figure (j) Difference in Histograms for previous technique with file size=32kb

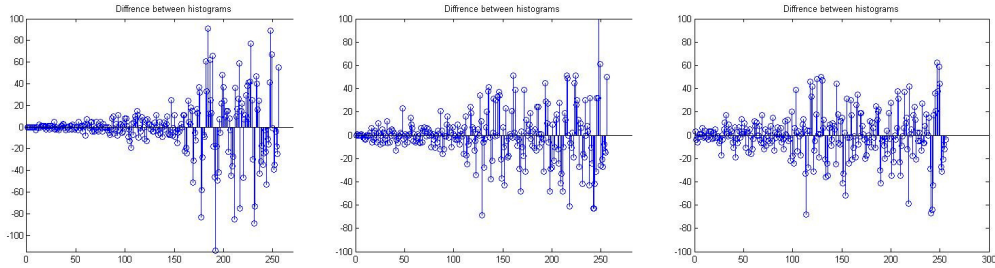


Figure 2(k) Difference in Histograms for proposed technique with file size =32 kb
 Figure 2. Comparison of Difference of Histograms

5.2.3. Impact on MSE and PSNR

These two parameters give an insight to check the variation on image quality mathematically. By analyzing the two parameters from Figure 3 following inferences can be made:

- As the file size increases the value of PSNR decreases and that of MSE increases. This is due to the fact that error increases by increasing the number of embedding pixels and PSNR is inversely proportional to MSE from equation (1) and equation (2).
- By using the proposed adaptive technique the changes in MSE and PSNR are marginal due to the fact that the length of data to be embedded is same.

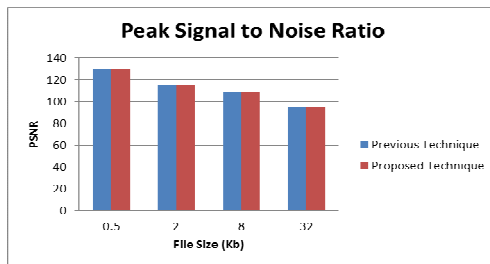


Figure 3(a) Comparison of PSNR

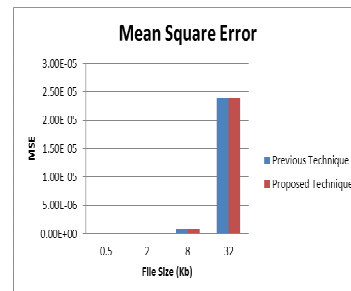


Figure 3(b) Comparison of MSE

Figure 3. Comparison of MSE and PSNR

5.2.4. Impact on time complexity

The time complexity of the proposed technique mainly depends upon the number of collision and the size of the data to be embedded. The following inferences can be drawn from the results (see Figure 4):

- Time complexity of proposed technique is less than that of the other technique due to the reduction in the number of collisions.
- As the file size increases the time complexity increases drastically.

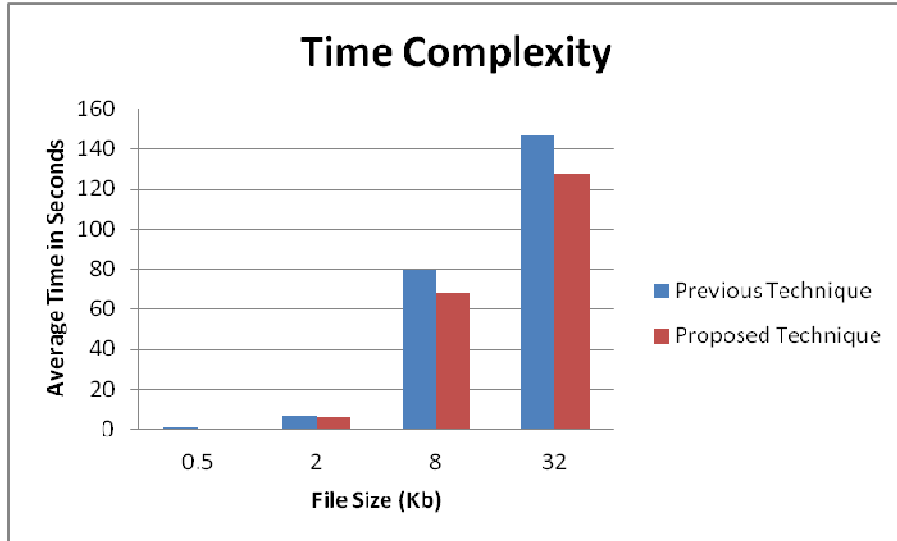


Figure 4. Comparison of time complexity

5.2.5. Collisions

The number of collision depends upon the size of cipher text to be embedded in image. From Figure 5, the following points can be inferred:

- The proposed technique reduces the number of collision in comparison to the previous strategy as the matter of fact that range of random pixel selection is adaptive in our case.
- The number of collision as expected to increase by increasing the file size.

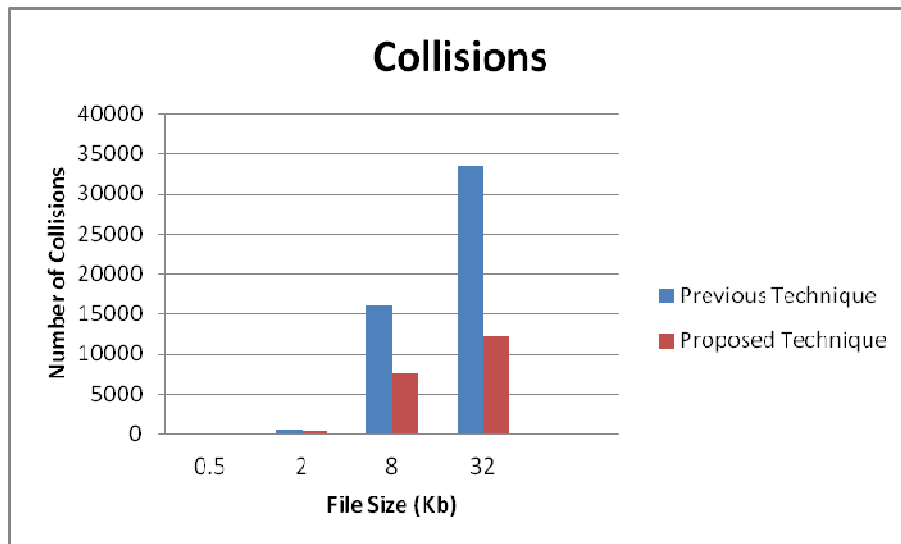


Figure 5. comparison of number of collisions

5.2.6. Distribution of Steganographic Noise

Steganographic Noise shows the distribution of the cipher text over the red, green and blue colour planes. From Figure 6 the following inferences can be drawn:

- The proposed technique distributes the cipher text evenly in all the planes. On the other hand the previous technique distributes high concentration in Red plane followed by Green and Blue plane.
- As the size of cipher text increases the distribution of cipher text crosses Red plane in case of previous technique.

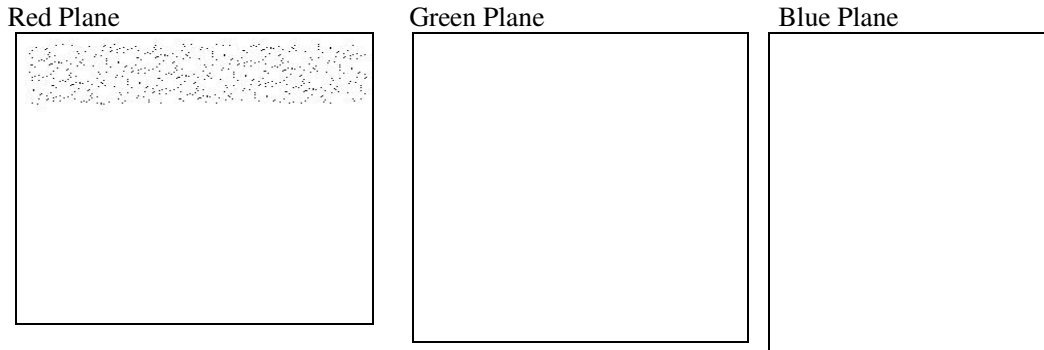


Figure 6(a) Distribution of steganographic noise for previous technique with file size =.5 kb

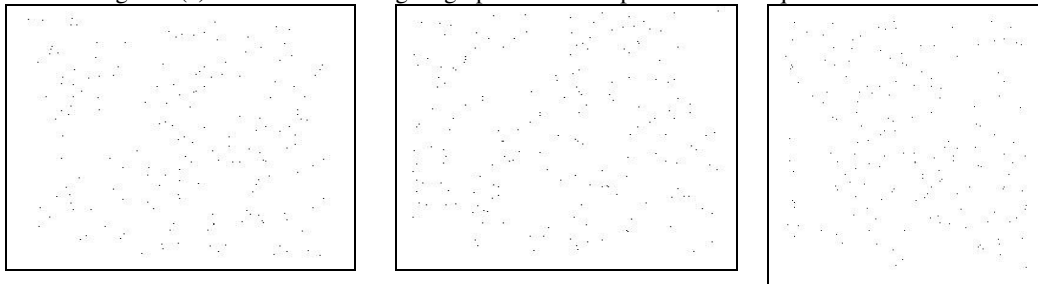


Figure 6(b) Distribution of steganographic noise for proposed technique with file size .5 kb

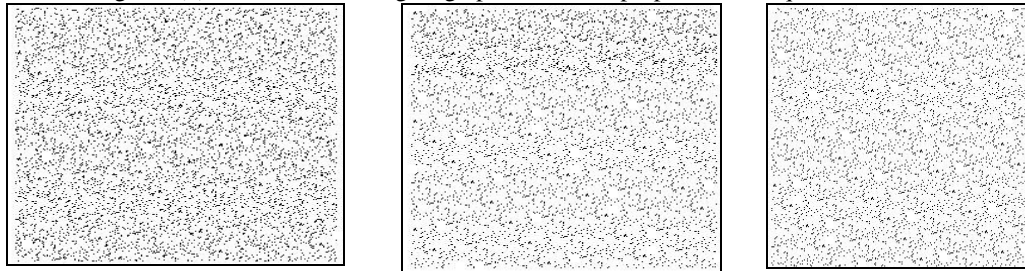


Figure 6(c) Distribution of steganographic noise for previous technique with file size=2 kb

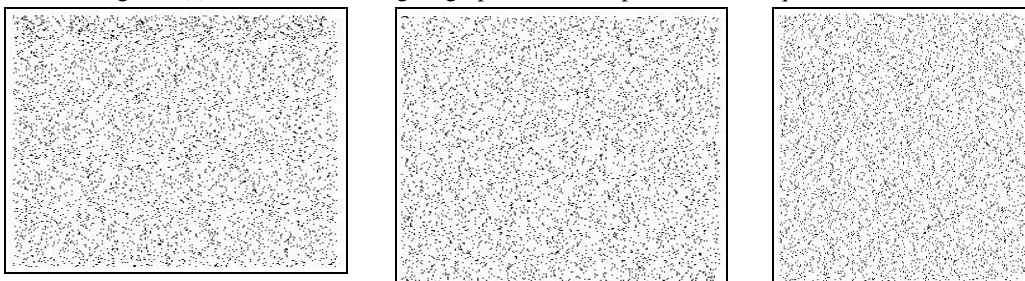


Figure 6(d) Distribution of steganographic noise for proposed technique with file size=2kb

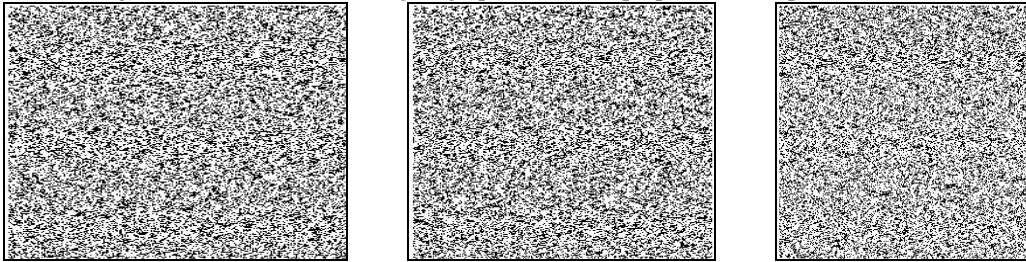


Figure 6(e) Distribution of steganographic noise for previous technique with file size=8 kb

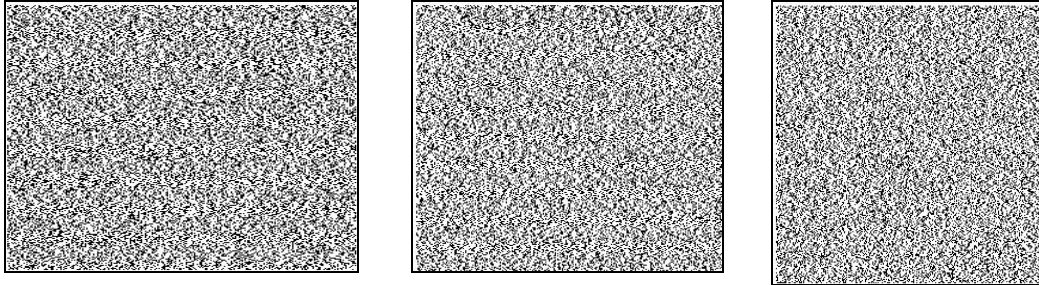


Figure 6(f) Distribution of steganographic noise for proposed technique with file size =8 kb

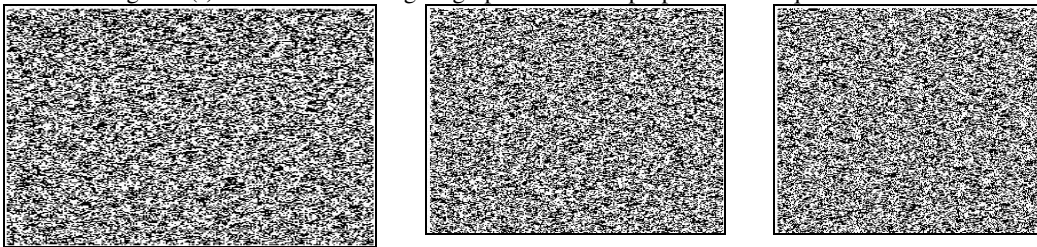


Figure 6(g) Distribution of steganographic noise for previous technique with file size =32 kb

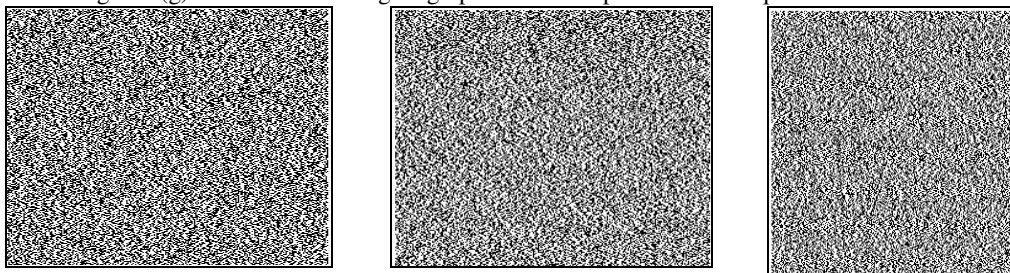


Figure 6(h) Distribution of steganographic noise for proposed technique with file size= 32 kb

Figure 6. Comparison of MSE and PSNR

5.2.7. Brute Force Search Time

In the previous technique message can be extracted with two keys: seed and decryption key whereas in the proposed technique third parameter as length of secret data also a mandatory requirement to extract the plaint text due to the fact that randomly selection of pixels also depends on the length of data. Despite having seed for pseudo random number generation an adversary cannot identify the embedded pixels. This provides additional security to secret data by increasing the brute force search time.

6. CONCLUSIONS AND FUTURE SCOPE

In this paper we propose an adaptive technique to provide enhanced security to plain text data by applying both cryptography and modified pseudorandom steganography. Following points can be inferred from the Table 3:

- Proposed technique provides more robustness against attacks by in terms of brute force search time. Also there is a need of Length requirement of data for an adversary to decrypt the plain text successfully.
- The number of collision in the proposed technique as compared to previous one is also very low implying the time complexity of the overall process to be low.
- The proposed adaptive technique evenly distributes the cipher text in all the colour planes in comparison to the previous one in which the distribution was not even.
- The value of MSE and PSNR of the proposed technique is nearly same in comparison to the previous one since the length of data to be embedded is nearly same.
- This paper shows how an adaptive technique can enhance the performance of stenographic process.

Table 3: Overall inferences of proposed scheme

Parameters	Previous Technique	Proposed technique
Brute Force Search Time	Low	High
Time Complexity	High	Low
Intensity of data distribution on colour planes	Vary at all the planes	Approximately same
Collisions	High	Low
MSE	Almost same	
PSNR	Almost same	

REFERENCES

- [1] "Steganography in Digital Media: Principles, Algorithms and Applications", Jessica Fridrich.
- [2] "Cryptography and Network Security principles and practices", William Stallings, pearsons education, first Indian reprint 2003
- [3] Kurak, C., and J. McHughes, "ACautionary Note On Image Downgrading," in IEEEComputer Security Applications Conference 1992, Proceedings, IEEE Press, 1992, pp.153-159.
- [4] Bender, W., D. Gruhl, and N. Morimoto, "Techniques for data hiding", IBM Systems Journal, vol. 35, no. 34, 1996, pp. 131-336.
- [5] M'oller, S., A. Pitzmann, and I. Stirand, "Computer Based Steganography How It Works and Why Therefore Any Restrictions on Cryptography Are Nonsense, At Best, in Information Hiding" First International Workshop, Proceedings, vol. 1174 of Lecture Notes in Computer Science, Springer, 1996, pp. 7-21.
- [6] Luby, M., and C. Racko, "How to Construct Pseudorandom Permutations from Pseudorandom Functions", SIAM Journal on Computation, vol. 17, no. 2, 1988, pp. 373-386.
- [7] Naor, M., and O. Reingold, "On the Construction of Pseudorandom Permutations" Luby-Racko Revisited, Journal of Cryptology, vol. 12, no. 1, 1999, pp. 29-66
- [8] J. Fridrich, M. Goljan, and R. Du, "Detecting LSB Steganography in Colour and Gray-Scale Images", Magazine of IEEE Multimedia Special Issue on Security, pp. 22-28, October-November 2001
- [9] Provos, N. "Defending Against Statistical Steganalysis". Proc. 10th USENIX Security Symposium. Washington, DC, 2001
- [10] J.J. Eggers, R. Bauml and B. Girod, "A communications approach to image steganography", Proceedings of SPIE, vol.4675, pp.26-37, 2002.

- [11] J. Fridrich, M. Goljan, "Practical Steganalysis of Digital Images – State of the Art", Proc. SPIE, Photonics West, Vol. 4675, Electronic Imaging 2002, Security and Watermarking of Multimedia Contents, San Jose, California, pp. 1-13, January, 2002.
- [12] J. Fridrich, M. Goljan, and T. Holotyak, "New Blind Steganalysis and its Implications", in Proc. SPIE Security, Steganography, and Watermarking of Multimedia Contents VIII, vol. 6072, pp. 607201, Jan. 2006
- [13] K. Solanki, K. Sullivan, U. Madhow, B. S. Manjunath, and S. Chandrasekaran "Provably secure steganography: Achieving zero K-L Divergence using statistical restoration" , Published in proceedings of ICIP 2006, pp. 125-128, IEEE 2006.
- [14] Shailender Gupta, Ankur Goyal, Bharat Bhushan, "Information Hiding Using Least Significant Bit Steganography and Cryptography", IJ.Modern Education and Computer Science, 2012, 6, 27-34
- [15] Rengarajan Amirtharajan, Jiaohua Qin, John Bosco Balaguru Rayapan, "Random Image Steganography and Steganalysis: Present Status and Future Directions", Information Technology Journal 11(5): 566-576, 2012
- [16] Shailender Gupta, Bharat Bhushan, Surbhi Singhanian and Jeetesh Gulani "A Hybrid approach for ensuring security in data communication", Accepted for publication in CCSIT 2013
- [17] Shailender Gupta, Bharat Bhushan, Surbhi Singhanian, Ajay Nain "A Novel Crypt-Stego Technique for Information Security in Communication Networks" Accepted for publication in IJSP 2013

Authors

Ajay Kumar Nain received his B. Tech degree from Department of Electronics and Communication Engineering at GJUS&T, Hisar, India in 2011. He is currently pursuing his M. Tech at YMCA University of Science and Technology, Faridabad, India. His current interests focus on image processing and Network security.



Shailender Gupta
Assistant Professor (Electronics Engg.)
YMCA University of Science and Technology,
Faridabad
E-mail: shailender81@gmail.com

Mr. Shailender Gupta is B. Tech (Electronics) and M. Tech and Ph. D (Computer Engg.) from YMCA University of Science and Technology. His academic interests include network security, automata theory and fuzzy logic.

Shailender Gupta
Assistant Professor (Electronics Engg.)
YMCA University of Science and Technology,
Faridabad
E-mail: shailender81@gmail.com

Mr. Bharat Bhushan has B. Tech (Electronics) from PEC and M.Tech (Electronics) from YMCA University of Science and Technology. His academic interests include Mobile Ad-hoc Network, Network Security.

Bharat Bhushan
Assistant Professor (Electronics Engg.)
YMCA University of Science and Technology,
Faridabad
E-mail: bhrts@yahoo.com

Ms. Rashmi Chawla has B. Tech (Electronics) and M.Tech (Electronics) from RGPVV, Bhopal. Her academic interests include Image Processing, Network Security and VLSI Designing.

Rashmi Chawla
Assistant Professor (Electronics Engg.)
YMCA University of Science and Technology,
Faridabad
E-mail: rashmi.chawla@rediffmail.com