Scientific Journal of Impact Factor (SJIF): 4.72

e-ISSN (O): 2348-4470
p-ISSN (P): 2348-6406

International Journal of Advance Engineering and Research Development

Volume 4, Issue 7, July -2017

HYBRID APPROACH FOR IMPROVING DATA SECURITY AND SIZE REDUCTION IN IMAGE STEGANOGRAPHY

Mandeep Kaur[1], Asst.Prof.Rupinder Kaur Randhawa[2]

[1]*Dept.of Computer Science and Engineering, Baba Banda Singh Bahadur Engineering College*
[2]*Dept. of Computer Science and Engineering, Baba Banda Singh Bahadur Engineering College*

Abstract — *Image steganography is used to hide the secret message under a cover image in order to enhance the security. There are several techniques proposed which has focused on the security issue in their work. In this paper, the proposed method uses three different techniques such as Huffman coding, DNA and ST. Initially, Huffman is applied over the text for the compression, and then DNA is applied over the compressed data for the encryption and lastly State Transition algorithm is used for updating the location in the image. The application of these algorithms is to provide high security in comparison with the traditional algorithms. Total three images will be used for the evaluation of traditional and proposed techniques where the message bit varies from 50 to 100. The parameters PSNR and MSE are used for the evaluation of their performance.*

Keywords - *Steganography,Cryptography,Huffman Encoding,DNA,State Transition.*

1. INTRODUCTION

In present day to day life, people exchange information with each other by means of various techniques such as cell phone or internet. But these techniques are not secure enough. As every person wants to keep their data about work to be secure and secret thus it is necessary to use such techniques that should provide security to both sender and receiver. Two techniques steganography and cryptography are helpful in the accomplishment of task for transferring and sharing of data in a secure manner. Sharing of data in concealed way is provided by these techniques.

For a secure data communication, steganography and cryptography are popular concurrent techniques that provide security against human interception by operating data due to cipher or cover their existence. Cryptography provides encryption approaches for protecting data without damage and secure. In case of Cryptography person can tell that a message has been encrypted, but he can't decrypt the message without using the appropriate key. In steganography, message itself may not be difficult to decrypt but person would not detect the existence of message. When these both techniques are merging together provides two levels of security. Steganography has benefit over cryptography in hiding of secret transmission, where in cryptography the visibility of the secret information attracts the attention of wire tapper.

Steganography is derived from the Greeks word "steganos" that defines covered writing. In ancient time, Greek used various methods to secure messages by writing a secret message on wooden tablets before concealing it with a fake writing on top of wax, or used to tattoo a message on a slave's head, then waiting for the hair growth for coverage then shave it back when it reaches the desired position. In present time, transmission of information in modern steganography systems is done secretly over public digital channels, within a bearer that appears to be nothing out of the normal.

1.1 Steganography

Steganography refers to the process of covering or protecting a file, message, image, audio or video within another file, message, image, audio or video. Steganography means securing one part of data within another part. Steganography is a technique that used to transfer secret information from a source to destination in such a way that a potential violator does not recognize the presence of the message.

The main target for the usage of Steganography is to keep away from drawing attention to the transmission of secret message. If intuition is raised, then this target that has been planned for achieving the security of the hidden messages, because if the hackers noticed any modification in the sent message then this viewer will try to know the secret information inside the message.

The main terminologies used in the steganography systems are as follows:

- Cover Message: Cover message is used to conceal messages, images into it.
- Secret Message: Secret message are the concealed materials in the steganographic process.
- Secret Key: Secret Key is used to embed the secret message on the basis of the hiding algorithm.
- Embedding Algorithm: Embedding algorithm is used to carry out the message concealing process.

In the Steganography system framework, before concealing process, the sender must select the suitable message bearer such as text, image, sound, video, and then select the effective hidden messages as well as the strong password and that must be known to the receiver. The productive and suitable Steganography algorithm must be selected that able to encrypt the message in more secure method. Then the sender may transmit the carried message with the secret information by chat or email, or by other advance techniques. Carried message with the secret information is the stego file. At receiver end, after receiving the message by the receiver, decrypt it with the help of extracting algorithm and uses the same password that is used by the sender.

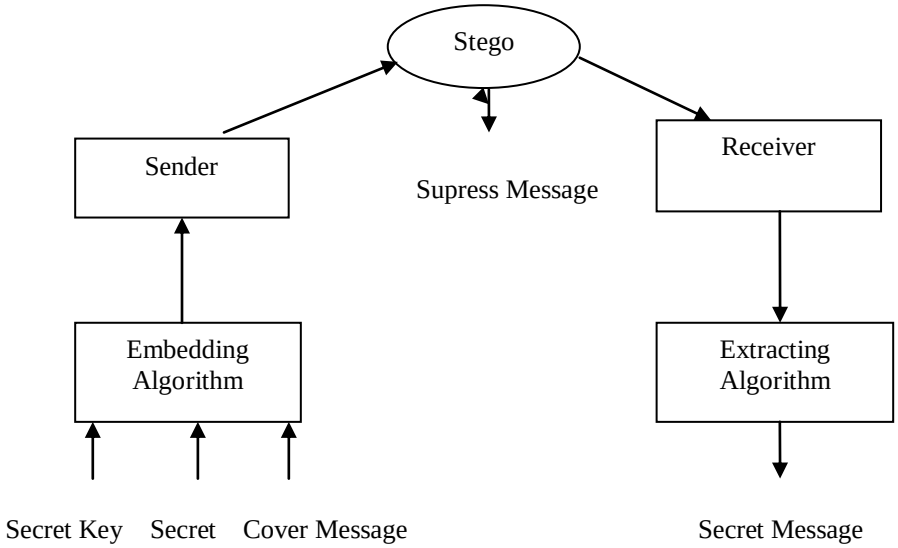

Fig.1- Steganography system scenario

In the recent technologies, various carrier messages can be used like Image, text, audio, video and many more. Most popular used for such purpose is the image file because it is very easy to transmit at the time of the transmission between the sender and receiver. Images are partition into the three categories: Binary (Black-White) images, Gray scale images and Red-Green-Blue (RGB) images. Binary image consist of 1 bit value per pixel (smallest pat of the screen) representing 0 for black pixels and 1 for the white pixels.

A Gray Scale image consists of 8 bits value per pixel representing 00000000 for the black pixels and 11111111 for the white pixels. While RGB (Red Green Blue) image consists of 24 bits values per pixel representing (00000000, 00000000, 00000000) and (11111111, 11111111, 11111111) for the white pixels. Red Green Blue (RGB) image is the most appropriate than other images because it contains large amount of data/information that will help in concealing the secret data/information with a bit variation in the image resolution (number of pixels in the image) and which doesn't lead to the distortion in the quality of the image and make the message more secured. In this paper, image used is the RGB (Red Green Blue) as a message carrier to conceal the secret message by the LSB (Least Significant Bit) hiding technique as well as the proposed technique.

1.2 Techniques for Steganography

Techniques for the steganography are classified such as:

- **Spatial Domain Technique**

Spatial Domain refers to add data to the image. Pixels of the randomly selected data are modified. Directly dealing with pixels of image is done in the spatial domain. In Spatial domain process data is hidden behind the pixels of image without applying any kind of encoding or the encryption method. Spatial domain uses LSB (Least Significant Bit) algorithm and many other.

i. Least Significant Bit (LSB) :

Least Significant Bit is one of the techniques of the spatial domain image steganography. Least Significant Bit is a technique for image steganography that works on the lowest significant bit in the byte value of the image pixels.LSB is a technique that doesn't affect the quality of image while data is embedded behind it.

Value of least significant bit changes but this change is not visible to the naked eye. Benefits of Least Significant Bit such as image doesn't distorted or mislead and by the use of LSB person can encode large amount of text

behind an image. LSB transmit the data to the receivers end with a great security and doesn't allow the intruder to access the encoded data.

- **Transform Domain Technique**

Various Frequencies are used in Transform domain technique for inserting the data in the image. For implementing the steganography it uses domain methods. Transform domain techniques are categorized as follow:

i. Discrete Fourier Transformation Technique (DFT):

Discrete Fourier Transform is a productive watermarking technique, the radiance of the frame that is watermarked will be produced and its magnitude of coefficients is taken to calculate DFT. Discrete Fourier Transform is a technique in which every pixel in spatial domain is transform to the frequency domain and it is partition into two parts real and the imaginary part.

ii. Discrete Cosine Transformation Technique (DCT):

DCT stands for Discrete Cosine Transform. The main characteristic of using DCT technique is that it will provide the good signal estimation by the use of certain coefficient values. Most of the algorithms use DCT for embedding the watermarking on the image.

DCT is an orthogonal transform (symmetric linear product) that make use of a basic functions with characteristics such as least bit error rate and large compression ratio. DCT break down the image into three bands namely low, middle and high and then it will make simplest to select the band in it the concealed data is to be embedded.

iii. Discrete Wavelet Transformation Technique (DWT):

In this technique of watermarking image is partition into four parts. These four parts are as diagonal part, vertical part, approximation part, horizontal part. For converting the image into low resolution, image will be partition into four parts. The process will be continued for computing the several scale wavelet decomposition.

DWT performs computations in accurate manner so it is a preferable technique for watermarking. The positive point of this technique is that it is strong to handle the noise in the image. Discrete Wavelet Transformation technique transforms the signals in time domain to the frequency domain. It is calculated by two filters such as low pass filter and high pass filter.

1.3 Methods of Steganography

Due to the advancement in the technology, security becomes a crucial factor. At the time of transmission, private data can be hacked by the intruder. Thus to secure the data there number of methods available. From which steganography is considered as successful methodology. The numerous methods of Steganography are:

- **Text Steganography**

In this type of Steganography text is hidden under the text files. This means that hidden data has been located at every nth letter of every words of text message. With the help of this method one can hide the data under cover text data. Some of the methods which can applicable are mentioned below:

- Format based Method
- Random and Statistical Method
- Linguistics Method

- **Image Steganography**

It is method of hiding the data under cover object referred as image. In image steganography image is a cover object for the hiding purpose. An image contains multiple pixels in RGB format thus intensity of the color is used to hide the data. Owing to this digital steganography has been used and digital images are used as cover source because of number of bits available in the digital representation of an image.

- **Audio Steganography**

A method which hides data in audio files is referred as audio steganography. There are number of formats available in Audio. Some of them are WAV, AU and MP3 sound file format. Numerous methods available in audio steganography are:

- Low bit Encoding
- Phase Coding
- Spread Spectrum

- **Video Steganography**

In video steganography data is hidden using digital video format file. Video is nothing but combination of images which are used as a carrier for embedding of secret data. In this steganography, DCT i.e. discrete cosine transform alter the values, these values has been used further for hiding of data in each images available in video. For example DCT changes the value 8.667 to 9 and further this value on the image will be considered for hiding of data. Furthermore, it is unnoticeable to the naked eye. Different formats available for video steganography are H.264, Mp4, MPEG, AVI etc.

- **Network or protocol Steganography**

In this type of steganography network protocol has used to hide the secret data. Some of the network protocol such as TCP, UDP, ICMP, and IP are considered as cover object. In the OSI layer of the network, there are some concealed channels available which provides steganography.

1.4 Types of Steganography

There are basically two types of Steganography named as fragile and robust. Below this explanation has been given for better understanding.

- **Fragile**

In this type of steganography, secret message can be destroyed if any changes or moderation have been made to the stegano file. For example if the file format of the stegano file is .bmp and if anyone changed its format into .jpeg or some other format so in such case hidden message or information will be ruined. Thus this type of steganography can be applied only to those where moderations is already made.

- **Robust**

In this type of steganography, information is not lost or destroyed as happens in fragile steganography which means that any moderation can be made after acquiring of stegano file. Robust steganography is difficult to implement in comparison with fragile.

1.5 Factors Affecting a Steganography Method

In the process of steganography, encryption and extraction algorithms performance can be checked by comparing stegano-image with the cover image. Various parameters can be considered in determining the efficiency of the technique or the algorithm used. Some of the factors are revealed below:

- **Robustness:** Robustness means that embedded data remain as it is even any moderation made to the stegano-image. In other words, after performing transformations on the stegano file, embedded data's meaning should remain constant. Transformation which can be performed on the stegano file are:

  i) Linear and non-linear filtering
  ii) Sharpening or blurring
  iii) Rotations and scaling
  iv) Cropping or Decimation
  v) Addition of random noise
  vi) Lossy Compression

- **Indistinguishable:** By the term indistinguishable, it means that how much steganographic algorithm provides invisibility to the data. As the concept of steganography is totally based on the invisibility of the secret data to the naked eye. Thus it should be the predominantly requirement of the algorithm. Moreover any steganography algorithm should provide and focus on this parameter.

- **Payload Capacity:** Payload Capacity is amount of the secret message or information that needs to be hidden in the cover object. Watermarking can also be used for hiding of data but it provides small amount of information.

- **PSNR:** PSNR is referred as Peak Signal to Noise Ratio. The ratio of signal should be high as compared to noisy signal which ensures the quality of the technique. If PSNR is high then the availability of the signal is higher and less noise which means technique is efficient. It compares original and compressed image to check the quality. Higher value ensures the quality of the compressed image.

- **MSE:** MSE means Mean Square Error which defines average difference between the referred image and distorted image. If MSE value is smaller then steganography technique is better and efficient contrary inefficient. This method is performed using pixel by pixel and adds squared difference of all the pixels and lastly divides total pixel count.

- **SNR:** Signal to noise ratio refers to ratio of signal and noise in the output. Comparison has done between the desired signals with the level of background noise.

1.6 Cryptography

Cryptography is a process which could be used to share data/information in an obscured manner. Cryptography includes moderation of a message in such a way which could be in digesting or encoded form secured by an encryption key which is known to both the sender and receiver only and without using secure/encryption key the message couldn't lead to access.

In the area of cryptography has a very important and rich history that has a range from pen and paper methods, to mathematical functions and especially to built machines that are in a use of today's routine. Coding and decoding secret messages is the process of cryptology. Cryptology is usually divided into cryptography that relates to design of cryptosystems for coding and decoding messages. It states that the term cryptography mainly refers to the combination of the cryptographic mechanisms that include:

- Digital signature schemes
- Encryption and decryption algorithms
- Integrity check functions

Cryptography is the process and study of methods for secure exchange information in the presence of third parties (called intruder). Mainly, it is all about to construct and analyze protocols that overcome the influence of intruders and which are related to several aspects in information security such as data integrity, data confidentiality and authentication.

- Plain Text: Plaintext is data/information a sender wishes to transfer to a receiver.
- Cipher text: Cipher text is obtained from the encryption that is performed on plaintext by the use of an algorithm, called a cipher.
- Encryption: It is the practice of transforming the image into some other image by the use of an algorithm so that any unauthorized person cannot access it. Only the person having an appropriate key can access that image.

1.7 Types of Cryptography

Two main types of cryptography are:

- Secret key cryptography
- Public key cryptography

Secret key cryptography (symmetric key cryptography): In this, when both the parties (sender, receiver) know the same secret code, called the key. Messages are encoded by the sender using the key and decoded by the receiver using the same key. In Secret key cryptography, key is shared between the sender and the receiver.

Public key cryptography (asymmetric key cryptography): Public Key Cryptography uses encryption and decryption algorithm pair. Sender and receiver make use of different keys for encryption and decryption.

DNA:

DNA stands for Deoxyribonucleic acid. DNA cryptography is one of the fastest emerging technologies that work on conception of DNA computing. DNA was developed in the 1994 by Leonard Max Adleman; development was done to solve the major problems like NP-complete problem similar to the Travelling Salesman problem, directed Hamilton path problem. 'A' in the RSA algorithm is also known for Adelman. The technique later on extended by various research

scholars for encoding and reduction in the data storage size that made the data communication over the network rapid and safe.

DNA is mainly used for the storing and the transmitting of the data. DNA Strands are long polymers of millions of linked nucleotides. These nucleotides consist of one of four nitrogen bases, a five carbon sugar and a phosphate group. The nucleotides that made up these polymers are named after the nitrogen base that it consists of; Adenine (A), Cytosine (C), Guanine (G) and Thymine (T).

2. LITERATURE SURVEY

R.S.A-EI et.al [13] describes discrete Cosine Transform based steganography uses DC parts to hide secret bits in a sequence manner in Least significant Bits. Their values indicate that the proposed tool provides a relatively a high embedding storage with no visual damage in the resulting image, whereas it enhances the security level and manages the correctness of the hidden data. S. Esra et.al [14] discussed about the storage and security problems of text steganography that have been considered to do better by propose scheme of a novel approach. The proposed scheme has been reached to 7.042% for hidden messages that contains 300 characters. Experimental results obtains that the proposed scheme provides a significant addition in terms of capacity. Q.T.Thach et.al [12] describes the work in need to find the modified pixels, or residuals, as an artifact of the embedding process. This paper establishes an excellence result that addresses shortcoming: shows that the expected mean residuals contain enough data to order logically the located payload provided that the size of the payload in each stego image varies. S.H.J et.al [15] discussed about three data hiding methods that are proposed, that are based upon the features of DNA sequences. For each method, a corresponding DNA sequence S is selected and the hidden message M is incorporated into it so that S' is produced. S' is then transfer to the receiver and the receiver is able to recognize and extract the message M hidden in S'. Finally, experimental results indicate a better performance of the proposed methods in comparison to the performance of the traditional methods.W.Kan et.al [19] describes technique is based on manipulating the quantization table and quantized discrete cosine transformation coefficients. Experimental reports show that the proposed method attains both high storage and high image quality without any damage. G.X.Z.D.Yang et.al [4] discussed about discrete STA, there are four basic operators like swap, shift, symmetry and substitute as well as the "risk and restore in probability" strategy. Firstly, main concern is on a parametric study of the restore probability p1 and risk probability p2. To effectively deal with the head pressure constraints, it investigates the effect of penalty coefficient and finds enforcement on the performance of the algorithm. Based on the experience gained from the training of the Two-Loop network problem, the discrete STA has successfully achieved the best known results for the Hanoi and New York problems. S.K. K et.al [16] describes paper proposes a scheme by compressing encoded data with the help of a subservient data and Huffman coding. The encoded data is then compressed using a quantization mechanism and Huffman coding. Experimental results show that the compression ratio distortion performance of this method is superior to the traditional Techniques. T. Turker et.al [18] discussed about a new data hiding technique is proposed based on hidden sharing scheme with the DNA exclusive operator for color images. The comparison of these techniques indicates that proposed techniques give the most successful result. M.Samiha et.al [11] describes paper presents a comprehensive analysis between the DNA-based play RSA, AES ciphers, fair, and vigenere each combined with a DNA hiding method. The conducted analysis results the performance diversity of each combined method in terms of security, speed, hidden storage in addition to both key size and data size. J.Reza et.al [6] discussed about, improvement of image compression through steganography. The first scheme in addition to a steganographic algorithm with the baseline DCT-based JPEG, while the next one uses this steganographic algorithm with the DWT-based JPEG. In this paper data compression is performed two times. Experimental results shows for this promising method to have wide potential /in image code.

3. METHODOLOGY AND BLOCK DIAGRAM

3.1 Methodology

Hiding the information/data in the image is defined as the steganography. Different types of techniques have been proposed but security of the information/data is one of the main reasons while the data is transmitted is not achieved. By studying different types of techniques a new technique is to be proposed in which the security level of the data is increased. Comparison by Huffman encoding technique will be applied to the data before the data is hidden. For encryption DNA is applied to the data. For hiding the data behind the image is done by the State Transition optimization. Security level will be increased by the encryption and compression of the data and it is not easy to detect. Proposed work aims to increase the security of the hidden data by applying HE, DNA, ST.

3.2 Block Diagram

Proposed work is partition into two categories. First one is for hiding the data and second one is for extracting the hidden data from stego image.

DATA EMBEDDING:

a) Cover image is selected and text is entered.
b) Huffman encoding technique is applied for compression so that size of data will be reduced and large amount of data will be hidden.
c) DNA is applied for encrypting the data so that security level will be increased.
d) State Transition is mainly used for hiding the data.
e) Finally, stego image is created.

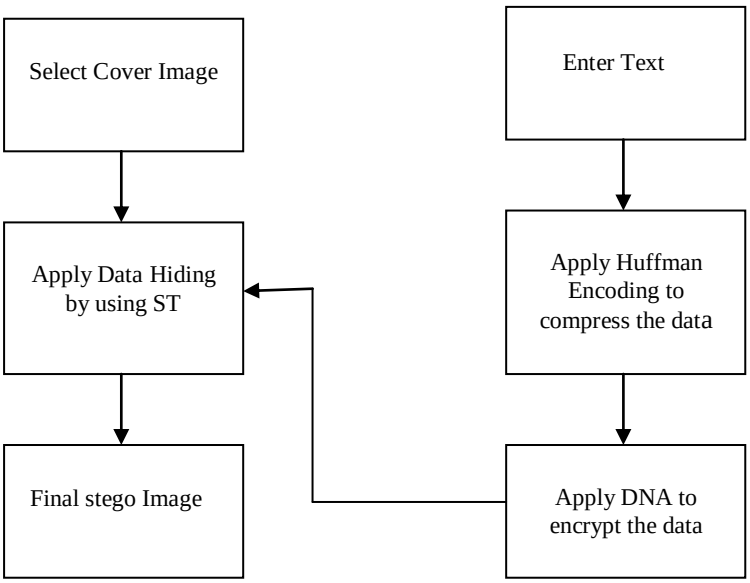

Fig. 2- Block diagram for data embedding

DATA EXTRACTION:

a) At receiver's side, stego image is selected.
b) Data extraction is performed on the selected stego image.
c) Decryption is done by the DNA.
d) Huffman algorithm performs the decompression.
e) After decompression the original text is extracted.

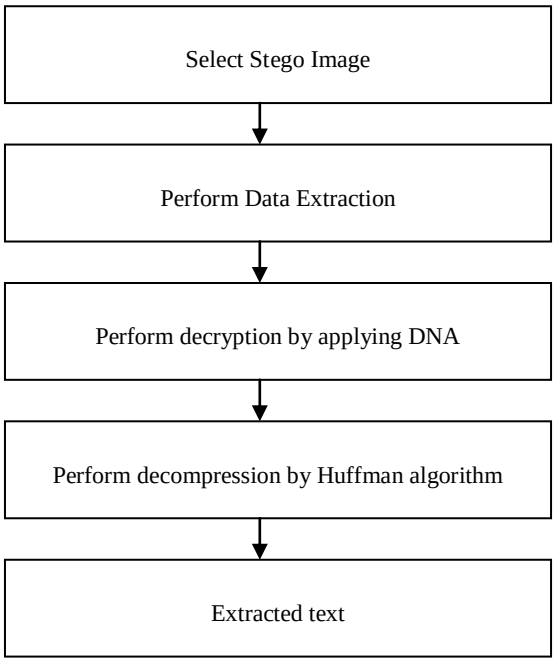

Fig. 3- Block diagram for data extraction

3.3 Parameters Used for Performance Evaluation

Parameters used for the performance evaluation are the PSNR (Peak Signal to Noise Ratio) and MSE (Mean Square Error).

3.3.1 PSNR

PSNR calculates the digital quality of an image. The ratio of signal should be high as compared to noisy signal which ensures the quality of the technique. If PSNR is high then the availability of the signal is higher and less noise which means technique is efficient. It compares original and compressed image to check the quality. Higher value ensures the quality of the compressed image i.e. better.

3.3.2 MSE

MSE calculates the average of the squares of the errors. If MSE value is small then steganography technique is better and efficient contrary inefficient. This method has performed using pixel by pixel and adds squared difference of all the pixels and lastly divides total pixel count.

4. CONCLUSION

The image steganography techniques have been used to embed the secret message under the cover image. Steganography provides better security for data sharing. DNA computing in the area of steganography and cryptography has been recognized as a possible technology that might bring forward a hope for unbreakable algorithms. The proposed technique will perform the similar task of steganography with two different algorithms such as DNA and Huffman. Whereas DNA algorithm is used for hiding and decryption of the text and Huffman is used for compression and decompression of the entered data. By compression and encryption of the data its security level is increased and it is not easy to recognize.

REFERENCES

[1] B.M.Baritha,V.Y," LSB Based Audio Steganography Based On Text Compression", in International Conference on Communication Technology and System Design, vol 30,Pp 703-710,2012.

[2] D.Prasenjit,D.Subhrajyoti,K.Nirmalaya,B.Baby," An Improved DNA based Dual Cover Steganography", in ICICT,vol 46,Pp 604-611,2016.

[3] E.N.N.EI, Z.R.A.S.AL,"New steganography algorithm to conceal a large amount of secret message using hybrid adaptive neural networks with modified adaptive genetic algorithm", in the Journal of systems and softwares, vol 86,Pp 1465-1481,2013.

[4] G.X.Z.D.Yang, S.A.R, "Optimal Design of Water Distribution Networks by Discrete State Transition Algorithm" Pp 1-14, 2013.

[5] J.I.F,D.K.A,Z.R.T.AI,S.R.R,"An efficient reversible data hiding algorithm using two steganographic images", in Signal Processing, vol 128,Pp 98-109,2016.

[6] J.Reza, Z.Djemel, R.M.Mehdi," Increasing image compression rate using steganography", in Experts Systems with Applications, vol 40, Pp 6918-6927, 2013.

[7] L.Hongjun, L.Da, K.Abdurahman,"A novel data hiding method based on deoxyribonucleic acid coding", in Computers and Electrical Engineering,,vol 39,Pp 11641173,2013.

[8] L.J.Fen,T.Y-Guo,H.Tao.Y.C-Fang,L.W-Bin," LSB steganographic payload location for JPEG-decompressed images", in Digital Signal Processing,Pp 1-11,2014.

[9] L.Y-Kai,"A data hiding scheme based upon DCT coefficient modification", in Computer Standards & Interfaces, vol 36,Pp 855-862,2014.

[10] M.P, K.T.Gireesh," Relating the embedding efficiency of LSB Steganography techniques in Spatial and Transform domains",in ICACC,vol 93,Pp 878-885,2016.

[11] M.Samiha,S. Ahmed, N. Khaled,"DNA-based cryptographic methods for data hiding in DNA media", in Biosystems, vol 150, Pp 110-118, 2016.

[12] Q.T.Thach,"Extracting hidden messages in steganographic images" in Digital Investigation,vol 11,Pp S40-S45, 2014.

[13] R.S.A-EI, "A comparative analysis of image steganography based on DCT algorithm and steganography tool to hide nuclear reactors confidential information" in Computers and Electrical Engineering, Pp 1-20, 2016.

[14]S.Esra,I. Hakan, "A compression-based text steganography method" in The Journal of Systems and Softwares, vol 85,Pp 2385-2394, 2012.

[15] S.H.J, Ng.K.L, F.J.F, L.R.C.T, H.C.H, "Data hiding methods based upon DNA sequences" in Information Sciences, vol 180,Pp 2196-2208, 2010.

[16]S.K.K,J.S.P,K.S.Kumar," Efficient Compression of Secured Images using Subservient Data and Huffman Coding" in RAEREST, vol 25,Pp 60-67,2016.

[17] S.K,P.P,T.Dr.P,M.Dr.C.M,"Dual Steganography approach for secure data communication",in International Conferencing on Modelling,Optimisation and Computing,vol 38,Pp 412-417,2012.

[18] T.Turker, A.Engin," A reversible data hiding algorithm based on probabilistic DNA-XOR secret sharing scheme for color images" in Displays, vol 41, Pp 1-8, 2016.

[19] W.Kan, L.Z-Ming, Hu.Y- Jian, "A high capacity lossless data hiding scheme for JPEG images" in the Journal of Systems and Software , Pp 1-11,2013.

[20] Z.Xiaojun, G.D.Yang, Y.Chunhua, G.Weihua," Discrete state transition algorithm for unconstrained integer optimization problems", in Neurocomputing, vol 173, Pp 864-874, 2016.

