Improved Steganography Method for Secured Data Sharing

Meetu Mann
Department of Computer Science & Engineering
MIET Meerut, India

Sudhir Goswami
Department of Computer Science & Engineering
MIET Meerut, India

ABSTRACT

Steganography is defined as the study of invisible communication. Steganography usually deals with the ways of hiding the existence of the communicated data in such a way that it remains confidential. It maintains secrecy between two communicating parties. In image Steganography, secrecy is achieved by embedding data into cover image and generating a stego-image. Everyday the development of internet communication is increasing in which the security of information is very important. Various techniques are used to hide data in different formats in steganography .Due to the simplicity of the least significant bit (LSB) substitution method, it is used to protected the data by converting data in digital image. This paper proposes a new data hiding technique for data to be secured which moves from sender to receiver. It gives more security in hiding data and more data can be secured via this technique. A stego image is used in it due to which a secure transmission of information is taken place without the distortion of the image. It is a secure way to keep the data confident. In this paper, we achieved a higher security than the previous work by using the multi key rather than the single key for decrypting the data in the Steganography image. We also compare the result in terms of PSNR, RMSE, MSE between the proposed work and previous work and got a better result analysis.

Keywords

Steganography, Encryption, Discrete Cosine Transform, Data Hiding, Security.

1. INTRODUCTION

In the present era, the communication is the basic requirement of each and field. Each of us wants to have in cognition of their data that is travelling over the digital medium. In daily routine life, many secure digital pathways are utilized like internet or telephone for transferring and sharing secret data, although we know that it's not secure completely. In order to share the information in a concealed manner two techniques could be used. The first method is cryptography and the second is Steganography [1] . In cryptography, the secret data is changed in a certain encrypted form by using encryption key .Only the sender knows what technique is used to encode the message this method is known as encryption key. The secret message cannot be decoded by anyone without using the encryption key. However, the transmission of encrypted message may attract attacker's doubt, and the encrypted message may thus be intercepted, attacked or decrypted violently. So, in order to overcome this difficulty of cryptographic techniques, steganography has been developed. Steganography is the art of communicating in such a way that it conceals the existence of the data. Thus, steganography hides the existence of information. In steganography the process of concealing information inside any multimedia content like digital picture, audio- video clip, is called "Embedding".

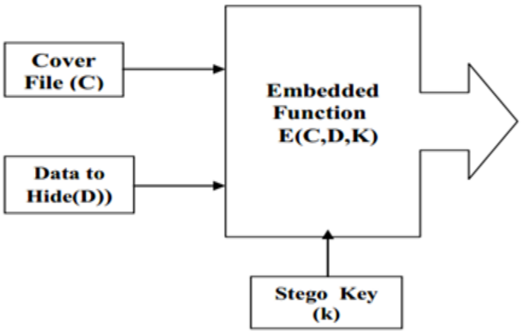


Figure 1: Steganography Working

1.1 Steganography Types

Steganography technique [2] can be divided as pure, symmetric and asymmetric. There is no need any exchange of information in pure steganography while, Symmetric Steganography and Asymmetric Steganography, does have a need to exchange of encoding technique in order to decode the data. Steganography is highly dependent on the type of media being used to hide the secret message. The medium that are commonly used are; text, digital images, audio files. Generally image Steganography is preferred. Advancement in technology of digital cameras and images that are being saved in cameras and then put these in computers has also enhanced many folds. Secondly, the secret text information hidden in the images does not destroy the picture the image and there are techniques that disturb only single bit of a digital image who's effects is almost negligible on its picture quality. But the major weakness of Steganography is that only small amount of data can be put behind in Steganography as secret message .Some methods are following.

(i) Encoding secret message in text /document.
(ii) Encoding secret message in audio clips.
(iii) Encoding secret message in digital images.

1.2 Text Steganography

Text Steganography embed [3] the secret information in text files with the help of several different methods :

(a) Format based Method
(b) Random and statistical method
(c) linguistic method

1.2.1 Format based Method

This method modifies the existing text for hiding the secret text. This technique involves the insertion of spaces, resizing the text, modifying the text format. In short we can say that this method, the secret message is hide by modifying the format of the existing text.

12.2 Random and statistical method

A random technique is used for hiding the characters that are appeared in random order. Statistical methods determine the statistics such as means, variance and chi square test which can measure the amount of repeated information to be hidden within the text.

1.2.2 Linguistic method

Linguistic method is a combination of syntax and semantics methods. Linguistic Steganography uses the linguistic properties of generated and changed text, and then uses linguistic arrangement as the space in which message is hidden. Syntactic Steganography analysis ensures that arrangements are syntactically correct. This is done because the text is produced by grammar, unless the grammar is syntactically flawed, the text it is guaranteed to be syntactically correct. In Semantic Method you can assign the value to synonyms and data can be encoded into actual words of text .

1.3 Audio Steganography

Hiding secret information behind a digital sound is known as audio Steganography. There are three techniques that are used in audio steganography are [4]:

1.31 Low Bit Encoding (LBE)

It is used in audio communications like mobile communications and VOIP. It executes, to embed the data while pitch period prediction is performed during low bit-rate speech encoding, thus keeping synchronization between information hiding and speech encoding.

1.3.2 Phase Encoding

This technique divides the actual audio file into parts and then the whole secret sequence into the phase spectrum of the very first block. One disadvantage of the Phase Encoding technique is that small message ability because message is kept in only first block.

1.3.4 Spread Spectrum Encoding

This method is a type of radio frequency communication. Data is sent using the spread spectrum encoding. One method of Spread Spectrum Encoding is DSSS (Direct Sequence Spread Spectrum). In this technique signal is spread by multiplying it by a fix maximal length pseudorandom sequence, it is also known as chip. Then after the calculation of start and end quanta is taken by the discrete, sampled nature of host signal for phase locking purpose. And as resultant, higher chip rate occur and we can hide maximum data in the chip.

1.4 Image Steganography

Digital images [5] are used as cover object Steganography. Image files are used for storing of digital images. An image file may store data in two formats these methods are compressed, uncompressed format. In Image Steganography, method for hiding data, can be classified into two categories. They are spatial domain and frequency domain. Spatial domain involve direct manipulation of pixels in a digital image. And frequency domain techniques is based on modifying the Fourier transform of a digital image. Steganography technique works on three types of images: Pallete based images (i.e.GIF images), Raw images (i.e ,BMP format) and JPEG images. One of the most popular format used on the internet is JPEG(Joint Photographic Expert Group). This method provides large compression ratio and maintains high image quality by measuring PSNR value.

In the JPEG compression, an image is divided into 8*8 blocks and then DCT is applied on each block. **Discrete Cosine Transformation** (DCT), is used for data compression. It is like Fast Fourier Transform, DCT converts data (pixels,waveforms, etc.) into sets of frequencies. After that resultant DCT coefficient matrix is quantized using a quantization table. Quantization table is a matrix, that contains DCT coefficients. In the end, inverse DCT of quantized coefficients is calculated and finally jpeg image is attained.

1.5 Application of Steganography

(i) Confidential Communication and Secret Data Storing
(ii) Protection of Data Alteration
(iii) Access Control System for Digital Content Distribution
(iv) E-Commerce
(v) Media
(vi) Database Systems.
(vii) digital watermarking.

2. BASIC THEORY

Steganography is the science of hiding information in data. Normally Steganography is finished intelligently such that it is difficult for an adversary to notice the existence of a hidden message in the otherwise innocuous data. The part of data that has the message embedded in it is visible to the world in the clear and appears as safe and normal. It is in stark contrast with cryptography where the message is scrambled to create it extremely hard or impractical for an adversary to put together. A message in cipher text arouses some sort of suspicion whereas invisible message embedded in clear text does not. This is the advantage of Steganography. The area of Information Hiding is concerned with concealing "secret" information of some type in "cover" information of another type. Secure methods for information hiding would have a great variety of both socially and commercially desirable applications. Here, two parties wish to conceal secret messages in "innocent-looking" communications over a public medium so that an eavesdropping adversary cannot detect the presence of the secret messages. We give a cryptographically-inspired, rigorous formalization of steganography, and prove that, for any channel, secure steganography is possible if and only if one-way functions exist [6].

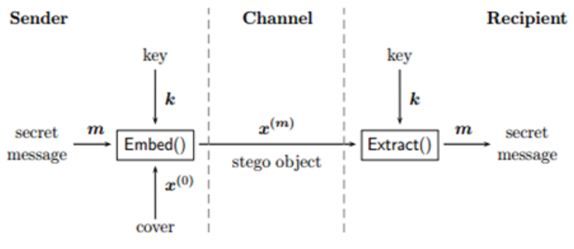


Fig 2 : Block Diagram of Baseline Steganography System

2.1 Image Steganography

Images are the most accepted cover objects used for Steganography. In the field of digital images many different image file formats exist, most of them for particular applications. For these different image file formats, different Steganography algorithms are available.

2.2 Image Definition

For a computer, a digital image is a collection of numbers that constitute different light intensities in different areas of the image. This numeric representation forms a grid and the individual points are referred to as pixels. Mostly, digital images on the Internet consists of a rectangular map of the image's pixels (represented as bits) where each pixel is located and its color. These pixels are showed horizontally row by row. The number of bits in a color scheme, called the bit depth, refers to the number of bits used for each pixel. The smallest bit depth in current color schemes is 8, meaning that there are 8 bits used to describe the color of each pixel. Monochrome and grayscale images use 8 bits for each pixel and are able to display 256 different colors of grey. Digital color images are normally stored in 24-bit files and use the RGB color model, also known as true color. All color variations for the pixels of a 24-bit image are resulting from 3 primary colors: red, green and blue, and each primary color is represented by 8 bits. Thus in one given pixel, there can be 256 different quantities of red, green and blue, adding up to more than 16-million combinations, resulting in more than 16-million colors. Not surprisingly the larger amount of colors that can be displayed, the larger the file size.

2.3 Image Compression

When working with larger images of greater bit depth, the digital images tend to become too big to transmit over a standard Internet connection. In order to show an image in a reasonable amount of time, techniques must be integrated to decrease the image's file size. These methods make use of mathematical formulas to analyze and condense image data, resulting in lesser file sizes. This procedure is called compression.

In images there are two types of compression: lossy and lossless. Both techniques save storage space, but the procedures that they implement differ. Lossy compression creates lesser files by discarding excess image data from the original image. It removes details that are too small for the human eye to differentiate, resulting in close approximations of the original image, although not an exact duplicate. An example of an image format that uses this compression technique is JPEG (Joint Photographic Experts Group). Lossless compression, on the other hand, never removes any information from the original image, but instead represents data in mathematical formulas. The original image's integrity is kept and the decompressed image output is bit-by-bit identical to the original image input. The most popular image formats that use lossless compression is GIF (Graphical Interchange Format) and 8-bit BMP (a Microsoft Windows bitmap file).

2.4 2D DCT (Discrete Cosine Transformation)

The 2-D DCT is a direct extension of the 1-D case and is given by.

$$C(u,v)=\alpha(u)\alpha(v)\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}f(x,y)$$

$$\cos\left[\frac{\pi(2x+1)u}{2N}\right]\cos\left[\frac{\pi(2y+1)v}{2N}\right]$$

for *u v N* , 0,1,2, , 1 = … - and *α*( ) *u* and *α*( ) *v* are defined in **(3)**. The inverse transform is defined as

$$f(x,y)=\sum_{u=0}^{N-1}\sum_{v=0}^{N-1}\alpha(u)\alpha(v)C(u,v)$$

$$\cos\left[\frac{\pi(2x+1)u}{2N}\right]\cos\left[\frac{\pi(2y+1)v}{2N}\right]$$

for *x y N* , 0,1,2, , 1 = … - . The 2-D basis functions can be generated by multiplying the horizontally oriented 1-D basis functions with vertically oriented set of same functions. The basis functions for *N* = 8 are shown in. Again, it can be noted that the basis functions exhibit a progressive increase in frequency both in the vertical and horizontal direction. The top left basis function of results from multiplication of the DC component with its transpose. Hence, this function assumes a constant value and is referred as DC coefficient.

2.5 Quantization

A highly beneficial property of JPEG method is that in this

step, varying levels of digital image compression and quality can be obtained by the selection of particular quantization metrics. This enables the user to decide on quality levels ranging from 1 to 100, where 1 is the lowest and 100 is the highest compression. As a result the quality compression ratio can be tailored to suit particular needs [7].

Subjective experiments involving human visual system have resulted in the JPEG standard quantization matrix. With quality level of 50, this matrix renders the both high compression and excellent decompressed image quality.

2.6 Image and Transform Domain

Image and Transform Domain Image steganography methods can be divided into 2 parts [8]: those in the Image Domain and those in the Transform Domain. Image – also known as spatial – domain methods embed messages in the intensity of the pixels directly, while for transform – also known as frequency – domain, images are first modified and then the message is embedded in the image.

Image domain methods encompass bit-wise techniques that apply bit insertion and noise manipulation and are sometimes thought as "simple systems". The image formats that are most suitable for image domain steganography are lossless and the method are typically dependent on the image format .

Steganography in the transform domain involves the manipulation of algorithms and image transforms. These techniques hide messages in more considerable areas of the cover image, making it more vigorous. Many transform domain methods are free of the image format and the embedded message may survive conversion between lossy and lossless compression.

In the next sections steganographic algorithms will be explained in categories according to image file formats and the domain in which they are performed.

2.6.1 Image Domain

- **Least Significant Bit**
- **LSB and Palette Based Images**

Least significant bit (LSB) [9] insertion is a frequent, simple approach to embedding information in a cover image. The least significant bit (in other words, the 8th bit) of some or all of the bytes inside an image is altered to a bit of the secret message. When using a 24-bit image, a bit of each of the red, green and blue color Text Images Audio/ video Protocol Transform Domain Image Domain JPEG LSB in BMP LSB in GIF Patchwork Spread Spectrum components can be used, since they are each represented by a byte. In other words, one can store 3 bits in each pixel. An 800 × 600 pixel image, can thus store a total amount of 1,440,000 bits or 180,000 bytes of embedded data.

LSB and Palette Based Images, for example GIF images, are another popular image file format commonly used on the Internet. By definition a GIF image cannot have a bit depth greater than 8, thus the maximum number of colors that a GIF can store is 256. GIF images are indexed images where the colors used in the image are stored in a palette, sometimes referred to as a color lookup table. Each pixel is represented as a single byte and the pixel data is an index to the color palette. The colors of the palette are typically ordered from the most used color to the least used colors to reduce lookup time.

3. PROPOSED METHODOLOGY

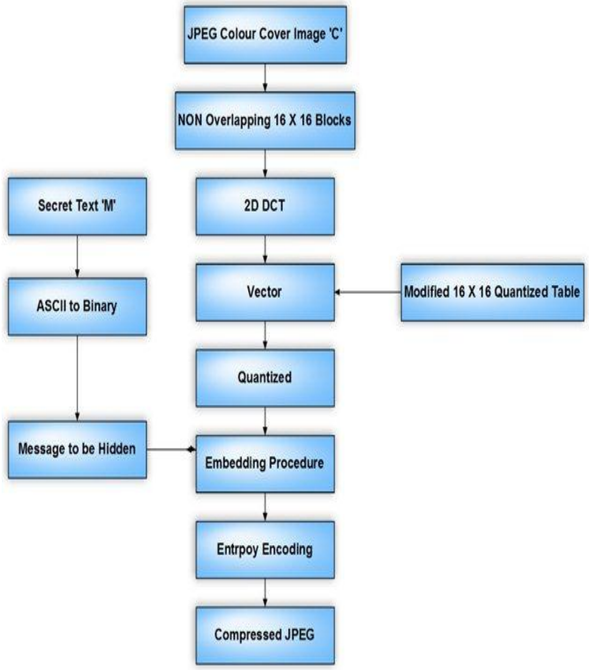


Figure 3: Embedding Procedure

The steps of proposed work -

- First of all, we have started process to upload the cover image which used to hide data.
- In the cover image pixel are break into Non overlapping16X16 blocks to hide the data.
- Show the dimension and size of original image.
- DCT is applied and update the quantization table.
- After that embedding procedure is applied to hide the data. Data is already converted in the binary form.
- Random key is generated and pass to the receiver for decryption.
- Entropy procedure is applied to hide the data.
- Data is hide or store in the LSB of image.
- Show the encoded image and show the dimension and size of image.

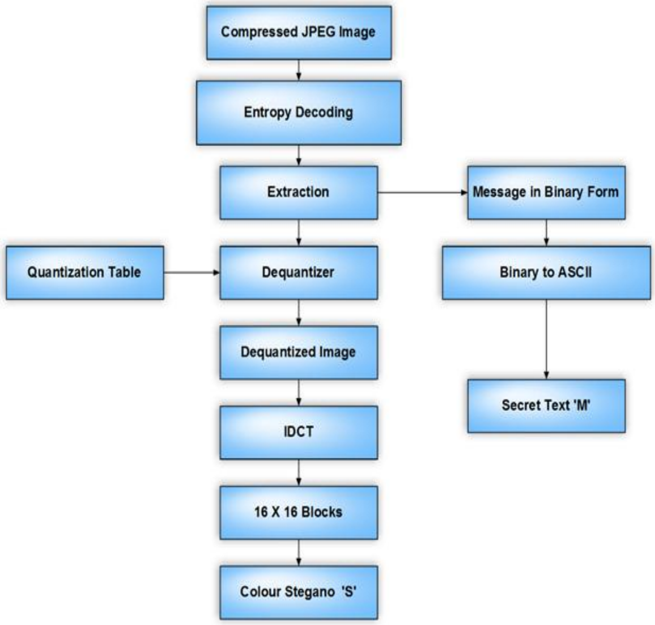


Figure 4: Extracting Procedure

4. EXPERIMENTAL RESULTS

By analyzing the below experimental results, it is obvious that the text hidden in the images can't be detected as the changes have been made only in the LSB, that does not distort the original image.

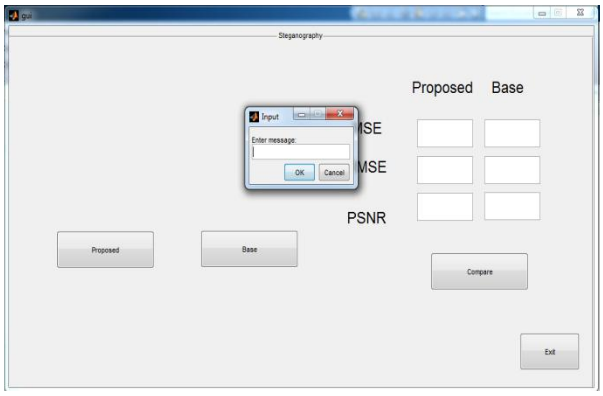
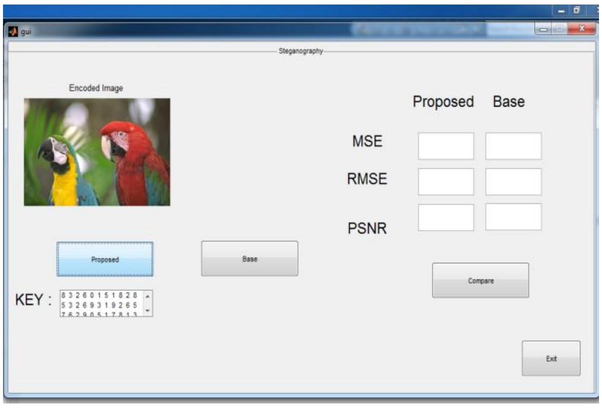
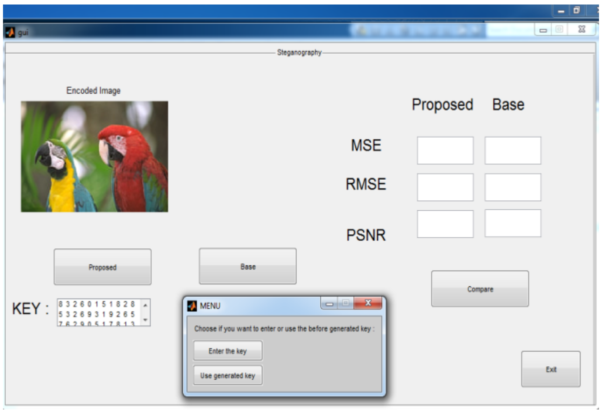


Figure 5: GUI for both code 16*16 DCT and 8*8 DCT for code execution. First show the input window to enter the message which want to hide in image using 16*16 matrix.



KEY : 8 3 2 6 0 1 5 1 8 2 8
5 3 2 6 9 3 1 9 2 6 5
7 6 2 9 0 5 1 7 8 1 3

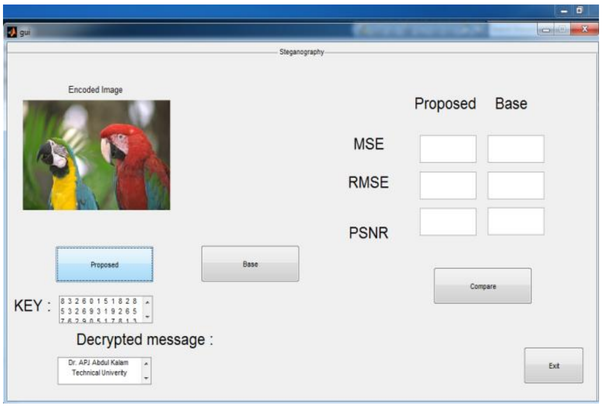
Figure 6: Show Encoded Image with Data. 16*16 DCT matrix is used to encode the data and random key is generated



KEY : 8 3 2 6 0 1 5 1 8 2 8
5 3 2 6 9 3 1 9 2 6 5
7 6 2 9 0 5 1 7 8 1 3

Choose if you want to enter or use the before generated key :

Figure 7: Show Encoded Image with Data. 16*16 DCT matrix is used to encod the data and random key is generated. And showing the message to decrypt the data with option Yes or No. After Yes, asking to enter key manually or use the random generated key to decrypt the message.



KEY : 8 3 2 6 0 1 5 1 8 2 8
5 3 2 6 9 3 1 9 2 6 5
7 6 2 9 0 5 1 7 8 1 3

Decrypted message :
Dr. APJ Abdul Kalam
Technical Univerty

Figure 8: After Yes, asking to enter key manually or use the random generated key to decrypt the message. After enter the key manually or random key message is decrypted and shown on window.

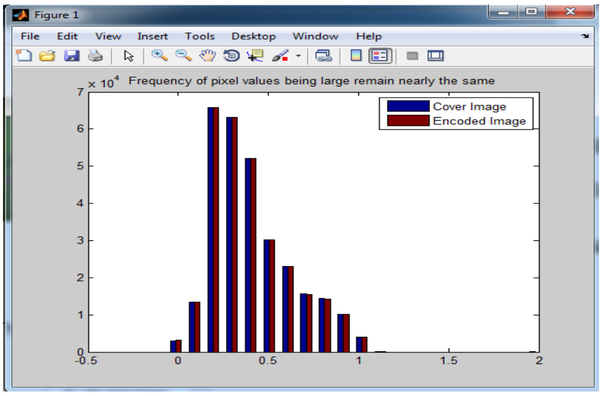
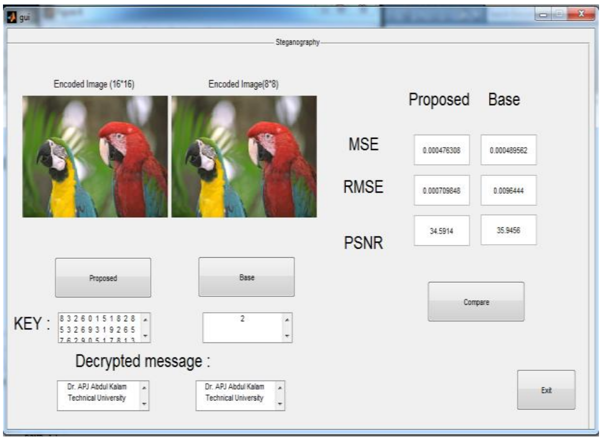


Figure 9: Histogram is generated for encoded is image and cover by using the 16*16 DCT matrix.



	Proposed	Base
MSE	0.000476308	0.000489562
RMSE	0.000709848	0.0096444
PSNR	34.5914	35.9456

KEY : 8 3 2 6 0 1 5 1 8 2 8
5 3 2 6 9 3 1 9 2 6 5
7 6 2 9 0 5 1 7 8 1 3

2

Decrypted message :
Dr. APJ Abdul Kalam
Technical University

Dr. APJ Abdul Kalam
Technical University

Figure 10 : PSNR, RMSE, MSE values are compare between proposed (16*16 DCT) and Base (8*8 DCT) works.

4.1 Result

As it can be seen in the Table that our technique is very efficient , it does not change the original image size, even after the secret encoded message.

Table 1: Performance Analysis

Technique	PSNR Value	RMSE Value	MSE Value	Key Size (For decryption)
16*16 DCT	45.6721	0.000709848	0.000476308	Long key
8*8 DCT	42.1748	0.0096444	0.000489562	Single key

5. CONCLUSION

In this work, it is evident that if we use Steganography we can share our secret data over a public medium without giving any doubt to the attacker. And the techniques we have used here in this is very efficient that it doesn't make any difference in the original and encoded image. Nobody can doubt over the image that has a secret data and as a future work more security features can be added with minimum changes in the cover image. To yield better imperceptibility the proposed method provided a higher similarity between the cover and stego pictures as a result. When steganography is combined with encryption a good security was achieved between two parties in case of secret communication, it is hardly attracted from eavesdropper by naked eye. Finally we can conclude that the proposed techniques effective for secret data communication and the size of the data to be hidden can be increased and also security level is increased. We also compare with it using the another method which 8*8 DCT and proved that the proposed work is efficient and highly secured to communication between two parties.

6. FUTURE WORK

Following are the various possibilities which can be done.

a. The proposed scheme can be compared with other techniques of Steganography.

b. In the proposed scheme pixel are breaks into 16X16 matrix which has high complexity and it is compared with 8X8 matrix. So, It can be replaced with some other matrix.

c. In future work increase the security level with hybrid techniques or by using the alphanumeric key

7. ACKNOWLEDGEMENT

My Humble and heartfelt acknowledgements to my esteemed guide **Mr. Sudhir Goswami** for his guidance and support without which this task would not have been accomplished. I would like to thank my friends and my colleagues, who have helped me during this research work. I am extremely grateful and indebted to my parents and my Colleague for being pillars of strength , for their unfailing moral support, and encouragement. I treasure their blessings and good wishes and dedicate this study to them. I thank one and all who have been instrumental in helping me to complete this dissertation work.

8. REFERENCES

[1] N. F. Johnson, S. Jajodia, "Exploring Steganography: Seeing the Unseen", IEEE Computer vol. 31, issue 2, pp. 26-34, 1998.

[2] J. C. Judge, "Steganography: Past, Present, Future", SANS Institute Publications, 2001.

[3] Artz D., "Digital Steganography: Hiding Data within Data", Internet Computing IEEE, vol. 5, issue 3, pp. 75-80, 2001.

[4] Jar no Mielikainen, "LSB Matching Revisited", Signal Processing letters, IEEE, vol. 13, issue 5, pp. 285-287, May 2006

[5] Anjali A. Shejul, Prof. U. L. Kulkarni, "A DWT based Approach for Steganography using Biometric", International Conference On Data Storage and Data Engineering, IEEE, pp. 39-43, 2010

[6] Priyanka Thakur, Santosh Kushwaha and Yogesh Rai. Article: Enhance Steganography Techniques: A Solution for Image Security. *International Journal of Computer* Applications 115(3):28-33, April 2015.

[7] Sudhir Goswami, Jyoti Goswami, Rajesh Mehra, IEEE International Conference on Recent Advances and Innovations in Engineering (ICRAIE-2014), May 09-11, 2014, Jaipur, India

[8] Kirti Shukla et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 6 (2) , 2015, 1632-1635

[9] G.Sateesh1 , E.Sai Lakshmi , M.Ramanamma3 , K.Jairam4 , A.Yeswanth, "Assured Data Communication Using Cryptography and Steganography", Volume V, Issue III, March 2016 IJLTEMAS ISSN 2278 – 2540