International Journal of Computer and Electrical Engineering, Vol. 2, No. 5, October, 2010
1793-8163

An Enhanced JPEG Steganography Scheme with Encryption Technique

Prosanta Gope, *Member IACSIT*, Anil Kumar and Gaurav Luthra

***Abstract*— In the aspect of information security, steganolysis has been an important topic since first indicated steganography has been used for communication. Apart from cryptography stegonagraphy is the additional method leadind to better secure of messages which goes hand by hand with cryptography, that's why revel of such a message not easy. A large number of JPEG steganography methods are available for free usage, which has spawned significant research in the area of JPEG steganolysis. This paper introduces an enhanced JPEG steganography along with a suitable encryption methodology which will play a significant role in the world of symmetric key cryptographic algorithm.**

***Index Terms*—Cryptography, DCT, MODDES, Quantization, Steganography.**

I. Introduction

Information security is and will continue to be a serious issue.Digital Steganography [6][8][10] has been one of the main weapon used to secure data. Secret information is imperceptibly hidden within signals with the use of steganography. It has already been proved that although cryptography is a measure concern as a information security [3][11] vehicle but if we merge steganography with a symmetric encryption technique like MODDES [1][2],X –MODDES , which will be beneficial in the aspect of security world .In the current world scenario we can not imagine all lives without computers but with usage a question of secure transfer of data appear very soon . Therefore coding, cryptography, Steganography is very important. Steganography [5] and cryptography are connected together more or less. Cryptography is strong in the usage of the key and the message is some how coded. But if we send an unsecured message, hacker [3] will notice it and will try to break it as well. But there is stegonagraphy, which helps with secure transfer of secret messages. It embeds a message inside a picture in such a way if we see the picture normally we can not recognize that there is is a secret message inside it.

This paper is then focused on the way how to embed messages inside the pictures and how to compress as well along with new encryption methodology has been introduced. Stegonagraphy methods can be classified on the basis of types of cover file which includes TCP/IP headers, images, and media data like video or audio file. This paper focuses images as cover, while this section gives a brief overview of existing steganographic methods.

Manuscript received October 9, 2001.

P.Gope is with the The Technological Institute of Textile and Sciences, Haryana, 127021 India. (Phone: +919728656745; fax: 01664-243728 ; e-mail: prosanta.nitdgp @ gmail.com).

Anil Kumar is with the The Technological Institute of Textile and Sciences, Haryana, 127021 India. (Phone: +919896343306; fax: 01664-243728; e-mail: yadavanil82@gmail.com

Gaurav Luthra is with the The Technological Institute of Textile and Sciences, Haryana, 127021 India. (Phone: +919253727292; fax: 01664-243728; e-mail: luthra.zenith@gmail.com

A. Transform based steganographic methods

It hides data in the coefficients of represented domain. In this technique we first map the signals to another domain such as Fourier transformation, DCT [14], Hartley transformation; the obtained co-efficient are altered.

B. Palette based Steganographic Methods

It hides the steganographic message within the bits of palette or indices. Care must be taken when using this image file format ensuring that number of colors is not exceeded.

C. Spatial Domain Based Steganographic Methods

The most commonly used of this type is LSB (Least Significant Bits) steganography in uncompressed file formats such as TIFF or BMP. The secret message may be hidden by altering least significant bit in a certain layer.

1) *Some basic Terminologies:*
   a) *Cover Image:* It is defined as original image into which the required information is embedded. We may also call it carrier image.
   b) *Stego Image:* It is an unified image obtained by combination of the cover image with payload.
   c) *Perceptibility:* It describes the ability of an intruder to visually detect the presence of hidden information in Stego image.
   d) *Robustness:* It is the ability of payload to survive the embedding and extraction process, even in face of manipulating of stego image such as filtering, cropping and rotating etc.

II. Error Analysis

Bit Error Rate (BER): For the successful recovery of hidden information the communication channel must be ideal but for real communication channel , there will be error while retrieving hidden information [7] and this is measured by BER.

$$BER = (1/|image^{cov}|) * \sum_{i=0}^{\text{all pixels}} |image^{cov} - image^{stego}|$$

Where cov = covered image
stego= Stego image

Mean Square Error (MSE): It is defined as the square of error between cover image and Stego image [12]. The distortion in the image can be measured using Mean Square Error (MSE).

$$MSE= \sum_{i=1}^{\text{all pixels}} \sum_{j=1}^{\text{all pixels}} [\{cov(i,j) - steg(i,j)\}^2 / NxN]$$

Where NxN is image size.

Peak Signal To Noise Ratio (PSNR): It is the ratio of the maximum signal to the noise in noise in Stego image.

$$PSNR = 20 \log_{10} (255/\sqrt{MSE})$$

.

III. Proposed Method

Step 1: The JPEG cover image [13] is broken down into 8x8 blocks of pixel.

Step 2: working from left to right top to bottom the DCT is applied to each block using the formula given below.

$$D(i,j)= (1/\sqrt{2N})\, c(i,j) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} p(x,y) \cos[\{2x+1\}i\Pi / 2N] \cos[\{2y+1\}i\Pi / 2N]$$

Where

$c(u)=1/\sqrt{2}$ if u=0

$= 1$ if u>0

Step 3: Find the corresponding matrix using the following equation

$$T_{i,j} = \frac{1}{\sqrt{N}} \quad \text{if } i=0$$

$$\sqrt{\frac{2}{N}} \cos\frac{(2j+1)i\pi}{2N} \quad \text{if } i>0$$

T=

.3536	.3536	.3536	.3536	.3536	.3536	.3536	.3536
.4904	.4175	.2778	.0975	-.0975	-.2778	-.4157	-.4904
.4619	.1913	-.1913	-.4619	-.4619	-.1913	.1913	.4619
.4157	-.0975	-.4904	-.2778	.2778	.4904	.0975	-.4157
.3536	-.3536	-.3536	.3536	.3536	-.3536	-.3536	.3536
.2728	-.4904	.0975	.4157	-.4157	-.0975	.4904	-.2728
.1913	-.4169	.4619	-.1913	-.1913	.4619	-.4619	.1913
.0975	-.2778	.4157	-.4904	.4904	-.4157	.2778	-.0975

The first row i =1 has all entries equal to $1/\sqrt{8}$ as expected from the formula.

Step 4: Now we apply Quantization and obtain a matrix of c's where

$C_{i,j} = \text{round}\ (D_{i,j}/Q_{i,j})$

Step 5: Now we are suggesting the encryption algorithm to be used:

Encryption Algorithm

Step I: Generate ASCII value of letter.

Step II: Generate corresponding binary value which should always be represented in 8 bit e.g. for decimal 12 should be 00001100 not 1100.

Step III: Reverse the 8 digit binary number.

Step IV: Take a 4 digit divisor therefore it should be greater than equal to binary 1000. It will serve as **key**.

Step V: Divide the reversed number obtained in step 3 by key.

Step VI: Store the remainder in first 3 digits and quotient in next 5 digits. If they are lesser than 3 and 5 digits respectively then add zeros to their left. This is our cipher text.

Decryption Algorithm

Step I: Multiply last 5 digits of cipher text by key.

Step II: Add first 3 digits of cipher text with the result obtained from step 1.

Step III: If the result produce in step 2 is not an 8 bit number we need to make it by appending zeros to its left.

Step IV: Reverse the number and find its corresponding ASCII value.

Advantages of this new Encryption Algorithm

1. There are two reverse operations present in this algorithm which would make it more secure.
2. CRC checking in receiving ends is easier
3. The algorithm is very simple in nature.

Disadvantage

Since it will work character by character therefore we need to apply the above algorithm for every character in our message.

Example

Suppose we have message "go" as this Encryption algorithm works character by character therefore we will first take the character g

Encryption

Step I: ASCII value of g is 103.

Step II: Binary value of decimal number 103 is 01100111

Step III: Reverse this binary number i.e. 11100110

Step IV: Now we will select 1000 as key.

Step V: Divide 11100110 by 1000

Quotient=11100 Remainder=110

Step VI: Therefore our cipher text would be 11011100. To ASCII (11011100)=220 which is "|(bar symbol)"

Hence our message "go" will convert into **"~"**

Decryption

Step I: On multiplying 11100 by key i.e 1000 we get 11100000.

Step II: Add first three digits 110 to the result obtained in step I. Therefore our result will be 11100110

Step III: Reverse the number obtained and we will get 01100111 which is 103 in decimal and it was our character 'g'.

Now we have

Step 6: Non-zero quantized DCT co-efficient of quantized matrix obtained in step 4 are used for purpose of finding the positions where our message characters would be added. If a_1, a_2, a_3, a_4 a_5, a_6, a_7 are non-zero quantized co-efficient and x_1, x_2 and x_3 are secret message which are encrypted by using our encryption algorithm then we will find positions b_1, b_2, b_3 where our message x_1, x_2 and x_3 are added as follows:

$b_1 = a_1$ XOR a_3 XOR a_5 XOR a_7

$b_2 = a_2$ XOR a_3 XOR a_6 XOR a_7

$b_3 = a_4$ XOR a_5 XOR a_6 XOR a_7

Where XOR represents bit by bit exclusive OR operation.

Selection of positions where our encrypted message will be added should be sequential and done according to following rule:

a. If b_i=0 then we will select from a $_{5,5}$, a $_{6,6}$,or a $_{7,7}$ as our position for x_1, x_2 and x_3
b. if b_i=1 then we may select from a $_{5,6}$, a $_{6,7}$ or a $_{7,8}$ as our position for x_1, x_2 and x_3

Step 7: As the quantized matrix obtained has non zero(positive or negative) entries on top left corner moreover these are very small in magnitude as compare to our message's value therefore we need to make these new entries Comparable with quantized matrix entries which is done as follows:

a. Convert our message character's ASCII value in string format.
b. Convert consecutive characters in decimal.
c. Use these decimal values for insertion in quantized matrix at specified locations.

Advantages

1. Encryption algorithm [2] makes our information safer.
2. We will substitute only one character of our encrypted message in that matrix Ci,j so that even if by steganalysis[6] anybody can determine that there is a slight change in image the information is very much distributed and of course encrypted .

Disadvantages

It may happen sometime that we need to revise $c_{i,j}$ matrix so there will be larger difference between Stego image and cover image.

IV. Case Study

Let we taken the following picture for our purpose

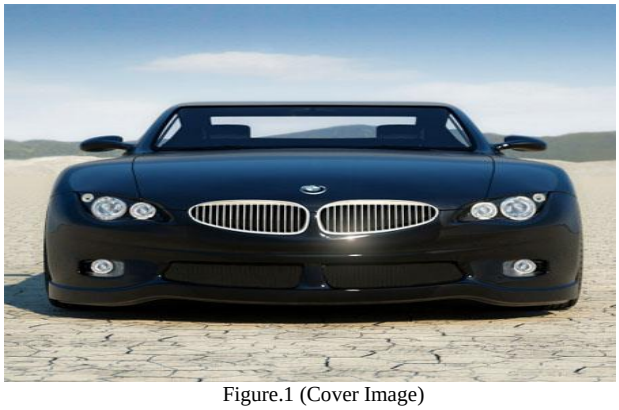

Figure.1 (Cover Image)

Therefore the corresponding color matrix of 8x8 blocks is given as:

Original=

154	123	123	123	123	123	123	136
192	180	136	154	154	154	136	110
254	198	154	154	180	154	123	123
239	180	136	180	180	166	123	123
180	154	136	167	166	149	136	136
128	136	123	136	154	180	198	154
123	105	110	149	136	136	180	166
110	136	123	123	123	136	154	136

Because DCT is designed to work on the pixel values ranging from -128 to 127 therefore we leveled off the original matrix by subtracting 128 from each entry i.e.

M= [Original] $_{i,j}$ -128

M=

26	-5	-5	-5	-5	-5	-5	8
64	52	8	26	26	26	8	-18
126	70	26	26	52	26	-5	-5
111	52	8	52	52	38	-5	-5
52	26	8	39	38	21	8	8
0	8	-5	8	26	52	70	26
-5	-23	-18	21	8	8	52	38
-18	8	-5	-5	-5	8	26	8

Now we will perform DCT which is accomplished by matrix multiplication given below:

D=TMT' Where T'= (adj T)/|T|

D=

162.3	40.6	20.0	72.3	30.3	12.5	-19.7	-11.5
30.5	108.4	10.5	32.3	27.7	-15.5	18.4	-2.0
-94.1	-60.1	12.3	-43.4	-31.3	6.1	-3.3	7.1
-38.6	-83.4	-5.4	-22.2	-13.5	15.5	-1.3	3.5
-31.3	17.9	-5.5	-12.4	14.3	-6.0	11.5	-6.0
-.09	-11.8	12.8	0.2	28.1	12.6	8.4	2.9
4.6	-2.4	12.2	6.6	-18.7	-12.8	7.7	12.0
-10.0	11.2	7.8	-16.3	21.5	0.0	5.9	10.7

This block now consist of 64 DCT coefficients, Ci,j where i and j range from 0-7 .The top left coefficient $C_{0,0}$ correlates

to low frequencies of original mage block where as $C_{7,7}$ corresponds to higher frequency . It is important that human eyes are most sensitive to low frequency.

Quantization

Q_{90} =

3	2	2	3	5	8	10	12
2	2	3	4	5	12	12	11
3	3	3	5	8	11	14	11
3	3	4	6	10	17	16	12
4	7	7	11	14	22	21	15
5	7	11	13	16	12	23	18
10	13	16	17	21	24	24	21
14	18	19	20	22	20	20	20

$C_{i,j}$=round($D_{i,j}$/ $Q_{i,j}$)

C=

10	4	2	5	1	0	0	0
3	9	1	2	1	0	0	0
-7	-5	1	-2	-1	0	0	0
-3	-5	0	-1	0	0	0	0
-2	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Figure.2 Stego image with Key and secured code sequence generated by X-MODDES

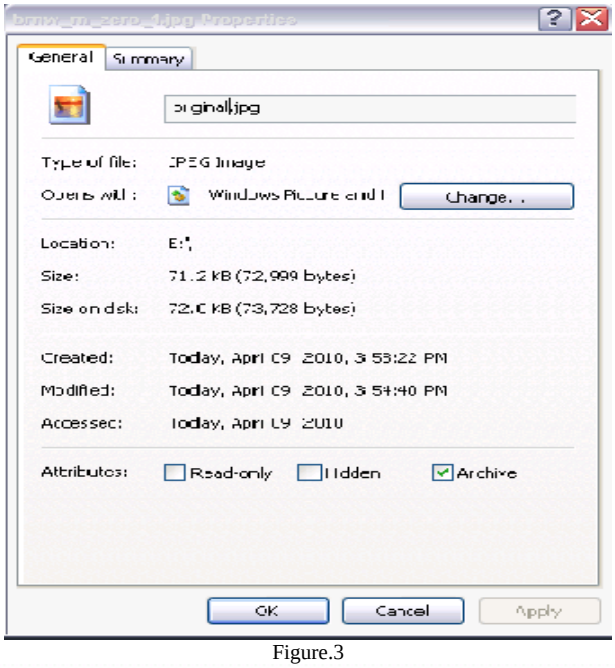

Figure.3

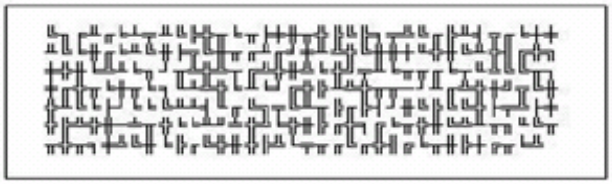

Figure 4 Secured Code Sequence of X-MODDES

Now after introducing the key value 144121 and the corresponding code sequence is on X-MODDES [4] we can see the changes in size of the file but the appearance of the image is same.

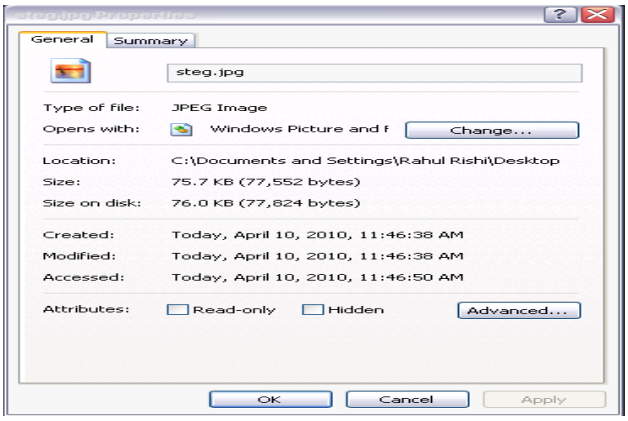

Figure 5

V. DECOMPRESSION

Reconstruction of our image begins by decoding the bit stream representing the quantized matrix C each element of C is then multiplied by corresponding element of the quantized matrix originally used .We will use the following formula

$R_{i,j}=Q_{i,j}$ X $C_{i,j}$

R=

$$\begin{bmatrix} 160 & 44 & 20 & 80 & 24 & 0 & 0 & 0 \\ 36 & 108 & 14 & 38 & 26 & 0 & 0 & 0 \\ -98 & -65 & 16 & -48 & -40 & 0 & 0 & 0 \\ -42 & -85 & 0 & -29 & 0 & 0 & 0 & 0 \\ 36 & 22 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

The IDCT is next applied to matrix are which is rounded to next integer and finally 128 is added which we subtracted at step 1 of compression process.

D= round (T'RT) +128

D=

$$\begin{bmatrix} 149 & 134 & 119 & 116 & 121 & 126 & 127 & 128 \\ 204 & 168 & 140 & 144 & 155 & 150 & 135 & 125 \\ 253 & 195 & 155 & 166 & 183 & 165 & 131 & 111 \\ 245 & 185 & 148 & 166 & 184 & 160 & 124 & 107 \\ 188 & 149 & 132 & 155 & 172 & 159 & 141 & 136 \\ 132 & 123 & 125 & 143 & 160 & 166 & 168 & 171 \\ 109 & 119 & 126 & 128 & 139 & 158 & 168 & 166 \\ 111 & 127 & 127 & 114 & 118 & 141 & 147 & 135 \end{bmatrix}$$

Comparison of matrices

Original =

$$\begin{bmatrix} 154 & 123 & 123 & 123 & 123 & 123 & 123 & 136 \\ 192 & 180 & 136 & 154 & 154 & 154 & 136 & 110 \\ 254 & 198 & 154 & 154 & 180 & 154 & 123 & 123 \\ 239 & 180 & 136 & 180 & 180 & 166 & 123 & 123 \\ 180 & 154 & 136 & 167 & 166 & 149 & 136 & 136 \\ 128 & 136 & 123 & 136 & 154 & 180 & 198 & 154 \\ 123 & 105 & 110 & 149 & 136 & 136 & 180 & 166 \\ 110 & 136 & 123 & 123 & 123 & 136 & 154 & 136 \end{bmatrix}$$

Decompressed =

$$\begin{bmatrix} 149 & 134 & 119 & 116 & 121 & 126 & 127 & 128 \\ 204 & 168 & 140 & 144 & 155 & 150 & 135 & 125 \\ 253 & 195 & 155 & 166 & 183 & 165 & 131 & 111 \\ 245 & 185 & 148 & 166 & 184 & 160 & 124 & 107 \\ 188 & 149 & 132 & 155 & 172 & 159 & 141 & 136 \\ 132 & 123 & 125 & 143 & 160 & 166 & 168 & 171 \\ 109 & 119 & 126 & 128 & 139 & 158 & 168 & 166 \\ 111 & 127 & 127 & 114 & 118 & 141 & 147 & 135 \end{bmatrix}$$

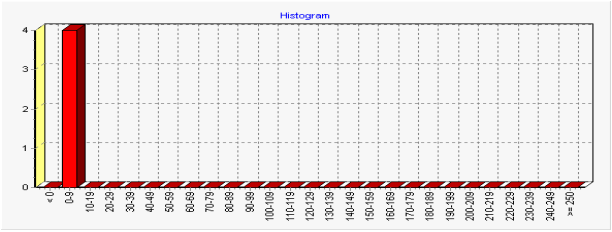

Figure7

Histogram of Cover Image (Figure.1)

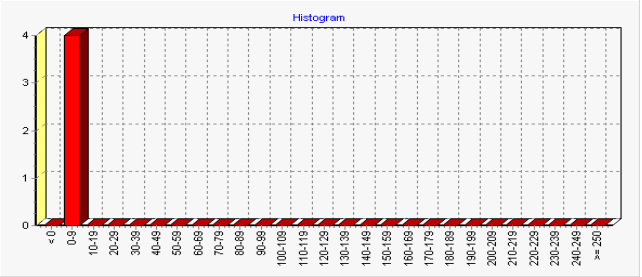

Figure 8

Histogram of Figure.2 (For the Image Embed With Message)

Now both the figure 7 and 8 clearly suggesting that even after embedding the message there is very slight change in histogram which is not observable and that's good from cracker prospective.

VI. Conclusion

In this paper we have proposed a new enhanced stegonagraphy methodology along with a suitable encryption scheme. As we can see there is a slight difference in the histogram therefore any smart hacker may find out that there is something secret in the image our encryption algorithm is enough for daunting his intention.

Acknowledgment

The authors would like to acknowledge and thank their parents, God and all the human being with great heart for their support and encouragement as well..

References

[1] P.Gope, "Multi Operator Delimiter based Encryption" (MODDES), ICCNT 2009.
[2] P.Gope, Dr.D.Gosh, "A Comparative study of Performance based Crypto analysis features for standard Data Encryption Algorithm with (MODDES), ICCNT, 2009.
[3] P.Gope ,"A New Secure Systematic Approach in the field of Network Security," ACVIT , 2010.
[4] P.Gope," Extended Multi Operator Delimiter Based Data Encryption Standard(X-MODDES)," ICFN, 2010, China.
[5] J.C.Judge, "Steganography: Past, present, Future," SANS whitepaper,November30,2001.
[6] Shuozhong Wang.Digital Steganography and steganalysis information war Technology in internet times, bejing: Tsinghua University ,Press ,2005,pp-70-72.
[7] Steganography: Hidden Data ,by Deborah Radcliff June 10,2002.http://www.computerworld.com/security topics/ security/story/0,10801,71726,00.html
[8] Bailey K,Curran K:Steganography(paperback).Book Surge publishing,2005,ISBN: 15945667X.
[9] Steganography Cited 2008-03-20 .http://en.wikipedia.org /wiki/steganography.
[10] D.Kahn,"The History of Steganography," Proceedings of the First International Workshop On Information Hiding, Lecture Notes in Computer science ,pp. 1-5,1996.
[11] G.R.Gordan, C.D Hosmer,C.Siedsma, and D.Rebovich," Assessing techenology, methods, and information for committing and combating cyber crime,"2003, http://www.ncjrs.gov/pdiffles1/nij/grants/198421.pdf.
[12] M.Kharrazi, H.T. Sencar, and N.Memon,"Performance study of common image steganography and steganalysis technique," Jouranal of electronic imaging,vol.15,no .4,app. 041104.1-041104.16,2006.
[13] J.Fridich,"features- based steganalysis for JPEG images and its implications for future design of steganographic schemes," in proceedings of the sixth international Workshop on Information Hiding, Lecture Notes in Computer Science,vol.3200. Springer verlag,2004,pp.67-81.
[14] T.Penvy and J.fridrich,"Merging marlov and DCT features for multi-class JPEG steganalysis," in proceedings of SPIE/IS&T,

Electronic Imaging, Security, Steganography, and Watermarking of multimedia contents IX, vol.6505,2007 ,pp.650 503.1-650 503.13.

Mr. Prosanta Gope become member of IACSIT in the year of 2008. He has received the master degree in computer science and Engineering from the "National Institute of Technology, Durgapur", India. Currently he is working as an Assistant Professor in the "The Technological Institute of Textile and Sciences", Haryana, India. He has his research contribution Five at International level in various proceeding like IEEE, WSP, Springer and one at National level. His research interest includes network security, cryptography and IP Routing.

Mr. Anil Kumar has received the master degree in computer science and Engineering from the "Guru Jambheshwar University of Science and Technology", Hisar, India Currently he is working as An Assistant Professor in the "The Technological Institute of Textile and Sciences", Bhiwani, Haryana, India. He has published 4 technical papers in carious national conferences and 1 in International conference. His research interest includes artificial neural network, wireless network and cryptography.

Mr. Gaurav Luthra is currently pursuing B. Tech. From the Technological Institute of Textile & Sciences, Haryana(India). He has his research contribution in two National levels. His research Interest includes Network Security, Ethical Hacking, Search Engine Optimization and Artificial Intelligence.

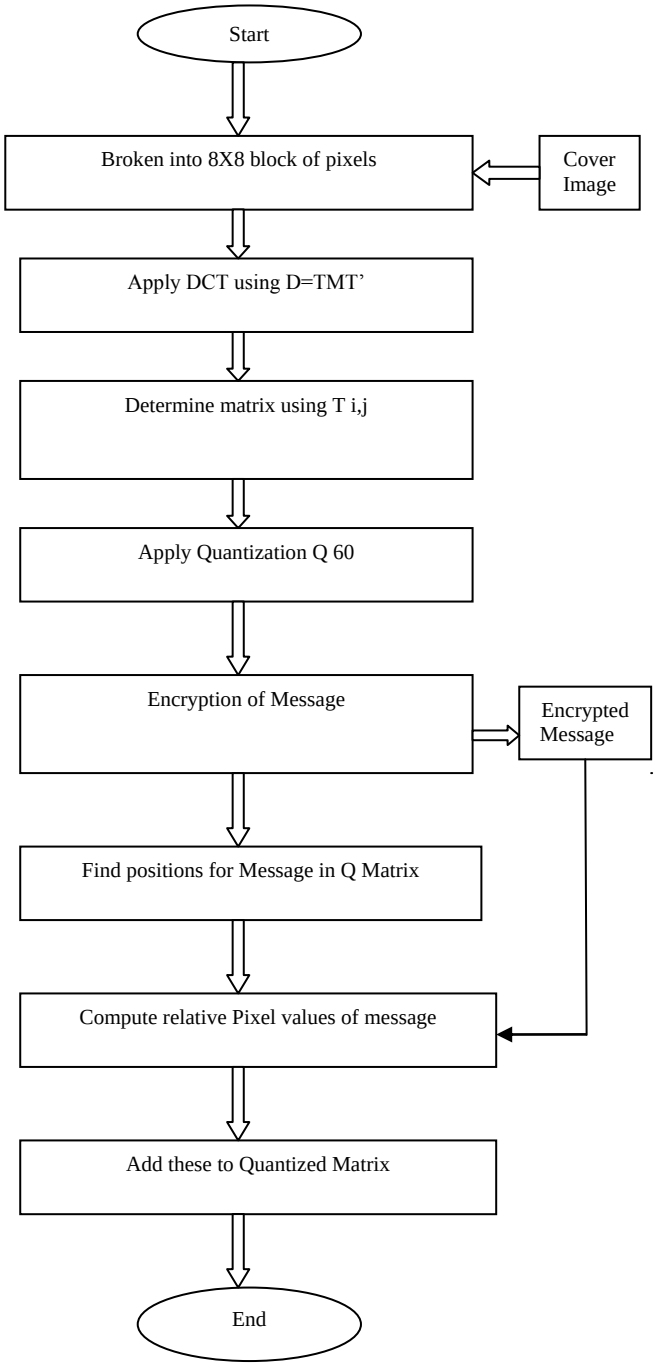

Figure 6 Flow chart of suggested method

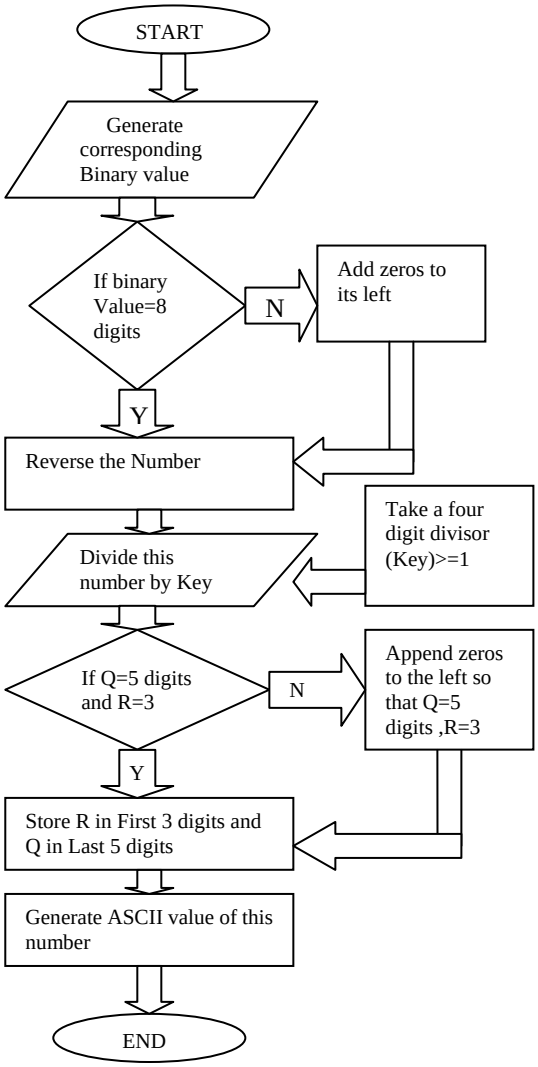

Figure 9 Flow chart of Encryption Process