International Journal of Advance Research in Science and Engineering
Volume No.06, Issue No. 12, December 2017
www.ijarse.com
IJARSE
ISSN: 2319-8354

Block Based Text Steganography

S.Thenmozhi[1], T.Gopala Krishnan [2], G.Santhana Mary[3]

[1]Associate professor, Dept of ECE, DayanandaSagar College of Engineering, (India)

[2]Assistant professor, Dept of ECE, Noorul Islam College of Engineering, (India)

[3]Assistant professor, Dept of ECE ,PSGIT, (India)

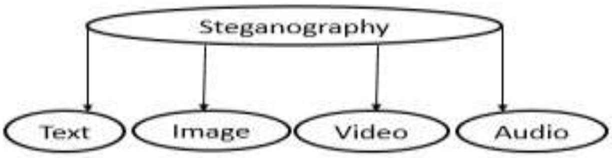
ABSTRACT

Profuse growth of multimedia communication through internet forces the security of data transmitted as an important feature. To guarantee the security of data communicated through unsecured guide, an image steganography is used. In Image steganography, secret information or image is veiled in the cover image so that the hackers or unauthorized intruders cannot notice. The paper proposes *a method which utilizes the pseudorandom number generator PRNG and Syndrome trellis code STC. PRNG is used to generate random number based on horizontal and vertical strip which act as key to generate random number and thus dynamic size of blocks are created in cover image. syndrome trellis code embed the secret bits in lsb bits of block image forming a route, the best route with minimum distortion and close to lsb bits of block image is selected. The STC bits chosen are combined with remaining bits of block image to create stego image. At the receiver side using horizontal, vertical strip as key random number is generated by PRNG and blocks are generated in stego image. Finally secret bits are extracted from block image. Proposed method enhances the security by using Block generation algorithm.*

Keywords– Steganography, Image Distortion, Pseudo Random Number Generator (PRNG), Syndrome Trellis Codes (STC).

I.INTRODUCTION

The promising growth of digitization technology brings into play of internet by which fast and easy data communication is possible. But along with the data transfer there is a security risk, personnel and official information can be hacked or robed by an intruder or an third party. Therefore to achieve secure communication many methods has been introduced among them Cryptography and steganography is widely used .Cryptography derived from Greek word Krypto’s meaning Hidden and graphein meaning to write. Cryptography is a method of modifying the secret message in such a way that it is unreadable for an unauthorized user or an attacker, Cryptography draws the suspicion of intruder towards the secret content. Steganography is a Greek word which means “concealed writing” “stegano” meaning “covered” or protected and “graphein” meaning to “write”. Steganography is categorized based on media it utilizes for hiding the secret data. and graphein meaning to “write”. Steganography is categorized based on media it utilizes for hiding the secret data.



Text Steganography: conceals the text behind some other text file. It is toughest type of steganography as the repetitive measure of text to hide the secret message is rare in text files.

Image Steganography: hides image in other cover media. It is the most frequently used strategy due to the restriction of the Human Eye.

Audio Steganography: is a method to transmit hidden data by adjusting a sound sign in an undetectable way. It is the science of concealing some secret text or audio data in a host message.

Video Steganography: Video Steganography is the procedure of concealing some secret data inside a video. The expansion of this data to the video is not conspicuous by the human eye as the change of a pixel color is negligible.

JyotiChaudhary, Neha[01] projected text steganography which utilizes the HTML and CSS (Cascading Style Sheets) coding of web pages as they are building Blocks of internet. Suspecting the presence of hidden text in web pages is impossible and security is further enhanced by hiding the Text in CSS coding of pages. Aruna Malik,Geetasikka,harshk.verna [02] proposed a Text Steganography Scheme based on LZW Compression and color coding. Proposed techniques uses forward mail platform to hide secret data. This algorithm first compress the secret data and then hides secret data in a cover .the secret data is embedded in a cover by making it colored using a color coding table. Swati Gupta and Deepti Gupta [03] discuss various types of text steganography techniques such as line shifting method, word shifting method, Syntactic method, Semantic method, New Synonym Text Method and among all these methods new synonym is better method to hide data. Tomas Filler et.al [04] proposed a practical approach for minimizing embedding impact in steganography based on syndrome coding and trellis-coded quantization it is based on syndrome coding using linear convolutional codes with the optimal binary quantize implemented using the Viterbi algorithm.

Wei-Wei Liu et.al [05] Syndrome-trellis codes (STCs) have been widely used in the field of data hiding. Above authors Proposed a scheme giving an alternative form of STCs based on quasi convolution codes, the parity-check matrix of which consists of single sub matrix while that of STCs contains two such ones..

II.METHODOLOGY

The proposed system architecture at the data embedding and extraction is represented in Figure 1 and Figure 2. Cover image can be chosen as bmp, jpg or png. A sequence of random numbers is generated by a PRNG (Pseudo Random Number Generation) with an arbitrary seed. Now the cover image can be divided into $H X W$ blocks i.e., horizontal strips and vertical blocks where, H denoted the number of strips and W as the number of blocks in each strip with various sizes using Raster scan manner.

On generating minimal route for the above divided blocks using STC (Syndrome Trellis Codes), the Secret text data which is converted into ASCII code and again ASCII code in to binary data is embedded into the LSB of the pixels along the route fashioned by Syndrome Trellis Codes. The path is stored as the best route and is used finally to create the stego-image. In the received stego-image the best path is extracted from the predefined region of the stego-image. Same as at the sending end, the stego-image is divided into C_h X C_v blocks using Raster scan manner. using LSB (Least Significant Bit) the data is extracted and conversion is performed on the binary format of stego-image into double and again binary to ASCII, to end with what we acquire is the secret data.

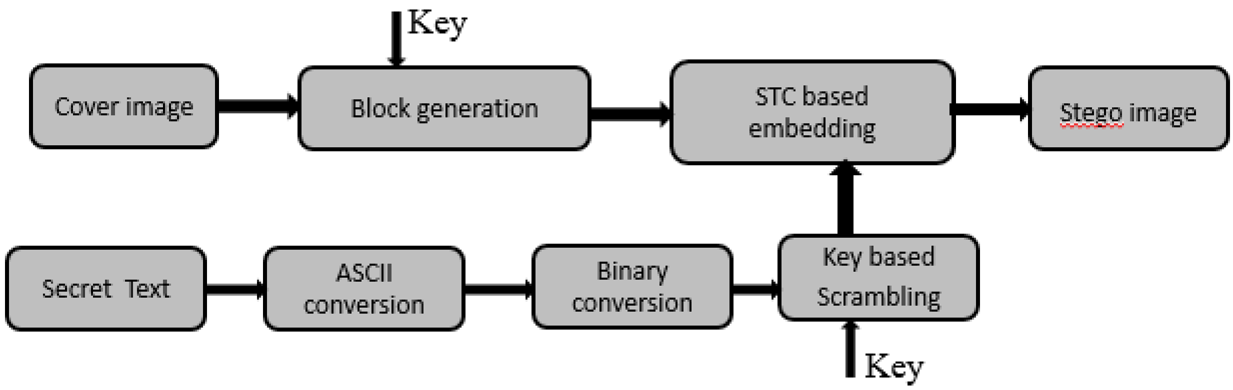


Fig.1. Block Diagram for Data Embedding

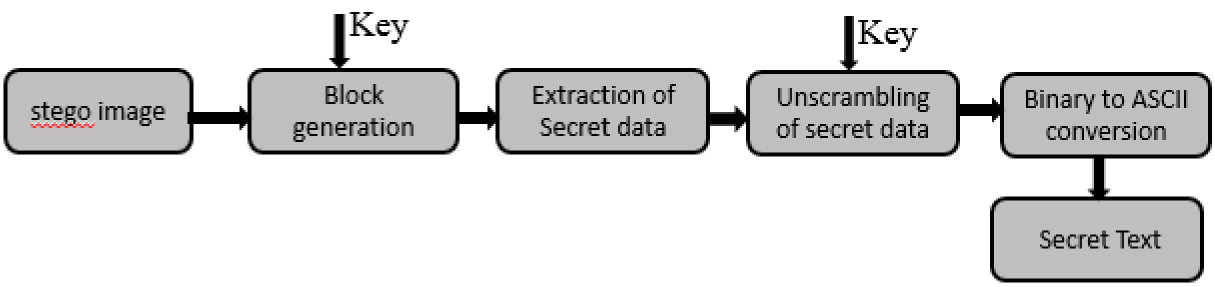


Fig.2.Block Diagram for Data Extraction

2.1 Generate Random Number

In the proposed method, the random number sequence is generated initially by making use of the Pseudo Random Number Generator (PRNG) with a seed which is arbitrary. Next stage is to divide the cover input image into the blocks of the dynamic size by making use of this sequence. These blocks are later traversed by using the manner of raster scan; the creation of the route is done. Secret information is embedded into the LSB pixels bit alongside with the route by making use of the Syndrome trellis codes. Te above procedure of creating route is repeated with different seeds, the route which leads to the minimal distortion is later chosen for the data embedding.

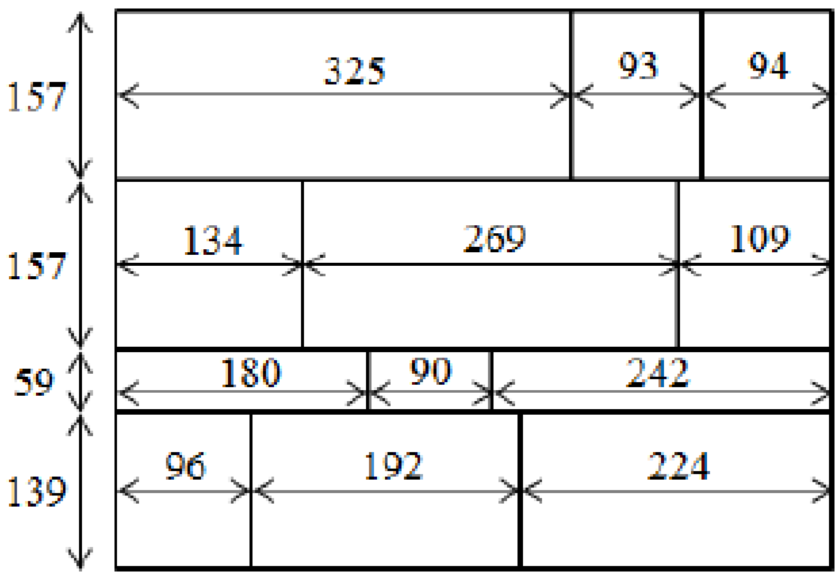
International Journal of Advance Research in Science and Engineering
Volume No.06, Issue No. 12, December 2017
www.ijarse.com
IJARSE
ISSN: 2319-8354

2.2 Blocking

This section gives the idea about how the cover image is divided into number of vertical and the horizontal blocks with varying sizes as depicted in the Figure 2 with the help of PRNG. In the proposed work the segment which is horizontal is called as the strip. Let us say that the c_h Will represent the strips number and c_v will represent the number of blocks present in each strips. For example let us consider that $c_h = 5$and $c_v = 4$. The strip width is always same as that of the cover image width and the block height is always equal to the strip height. Hence it is very necessary to have the strip height c_h and also the weight of the blocks ($c_h x c_v$). These sizes are yielded by making use of the sequence of different numbers) $N_i (1 \leq i \leq c_h + c_h \times c_v)$ which is generated from the PRNG. Let $|S_i|$ $(1 \leq i \leq c_h)$ is the strip height which is computed using the eq. (01),

$$|S_i| = \begin{cases} \left\lfloor \frac{N_i}{\sum_{j=1}^{c_h} N_j} \times h \right\rfloor & \text{where } 1 \leq i \leq c_h \\ h - \sum_{j=1}^{i-1} |S_j| & \text{where } i = c_h \end{cases} \qquad (01)$$

Where h represents the cover image height. Once the strip height is determined, the block width of every strip is obtained by utilizing the number sequence. Let $|B_{ij}|$ $(1 \leq i \leq c_h)$ is the Block j width present in the strip i which is calculated with eq. (02),



$$|B_{ij}| = \begin{cases} \left\lfloor \frac{N_{k+j} \times w}{\sum_{l=k+1}^{k+c_v} N_l} \right\rfloor \text{where} & \begin{matrix} 1 \leq i \leq c_v \\ k = (i-1) \times c_v + c_h \end{matrix} \\ w - \sum_{k=1}^{j-1} |B_{ik}| & \text{where } j = c_v \end{cases} \qquad (02)$$

The blocking procedure can be explained well with an example discussed in this section. If we consider $c_h = 4$ and also $c_v = 3$. It is very likely to have $4 \times (3 + 1) = 16$ number of random numbers. Let us assume that the number sequence which is obtained by the PRNG at this instant be 8,8,3,7,7,2,2,5,10,4,6,3,8,3,6,7. Considering this the four strips height in the cover image of size 512x512 is obtained using the eq. (03),

$$|S_1| = \left\lfloor \frac{8}{8+8+3+7} \times 512 \right\rfloor = 157$$

$$|S_2| = \left\lfloor \frac{8}{8+8+3+7} \times 512 \right\rfloor = 157$$

$$|S_3| = \left\lfloor \frac{3}{8+8+3+7} \times 512 \right\rfloor = 59$$

$$|S_4| = 512 - (157 + 157 + 59) = 139 \qquad (03)$$

We engaged the first four numbers of the sequence for determining the height of the strips. The remaining twelve numbers will be used for obtaining the width of three blocks in each of four strips [07], [08], [09]. These sizes are depicted in Fig. 3. For example, the width of three blocks of strip 1 was computed by eq. (02) as given below,

$$|B_{11}| = \left\lfloor \frac{7}{7+2+2} \times 512 \right\rfloor = 325 \qquad (04)$$

$$|B_{12}| = \left\lfloor \frac{2}{7+2+2} \times 512 \right\rfloor = 93$$

$$|B_{13}| = 512 - (325 + 93) = 94$$

2.3 Syndrome Trellis Code

In this section we are paying attention more on solving the problem of the binary PLS with the distortion and also modify the standard trellis coding strategy for the steganography. The code which is resultant called as the syndrome trellis codes. These codes will always serve as the block of building for the non-binary problems. The construction in attendance behind the STCs is also not new from the perspective of the information theoretic, since these STC's are the convolution codes represented in the dual domain. In case of practical steganography STC's are very interesting, as they help in solving both the problems of the embedding with a coding loss which is very small over a profiles of distortion even in case of the wet pixels. Always STCs will offer the solution for the problems in the embedding for the steganography. Descriptions for the codes along with their graphical form of the representation are given. All of these constructions are prepared with the Viterbi algorithm which is very optimal in case of problem solving. The idea behind the Syndrome Trellis Codes is to assign the embedding distortion to each of the cover element and also to embed the payload with the necessary distortion possible. Equation (05) defines the extraction and embedding mapping as,

$$Embedded: \{0,1\}^n \times \{0,1\}^k \rightarrow \{0,1\}^n$$

$$Extraction: \{0,1\}^n \rightarrow \{0,1\}^k$$

$$Embedding(x, m) = y,$$

$$Extraction(y) = m$$

$$\forall x, y \in \{0,1\}^n, \forall m \in \{0,1\}^k \qquad (05)$$

Where x denote the vector of cover, m denotes the sequence of the message and y denotes the vector of the stego. These methods can find the modulation of the k bit message and the cover of n element, while just keeping the distortion which is expected as small as possible. In this syndrome trellis coding, the extraction and the embedding mapping will be realized by using the linear code C which is binary of the n length and also dimension n - k. By assuming the parity check matrix H in the eq. (01), the mapping of the extraction is as given in the eq. (06),

$$Extraction(y) = H_y = m \qquad (06)$$

By Assuming as in eq. (07),

$$C(m) = \{z \in \{0,1\}^n\} | H_z = m \qquad (07)$$

The above given equation is the coefficient set which is corresponding to the sequence of the message m, these STC will also find the z optimal, which is closer to the x from the coefficient and take the output as y as given in the eq. (08)

$$y = Embedded(x, m) = arg \min_{y \in c(m)} (D(x, y)) \ (08)$$

It is seen that the STC will improve the efficiency of the embedding very effectively at the embedding rate which is low, i.e. when $\alpha < 0.5.$ the overall performance will always decrease at the increased embedding rate i.e. $\alpha \geq 0.5$).

2.3 Data Embedding Technique

Initially c_h, c_v and the determination of the n is done arbitrarily. If the repetition of the iterative loop is done for n times in the initial step, the generation of the random number is done as a seed and the generation of the sequence $c_h \times (c_{v+1})$ number generation is done by making use of the PRNG making use of the seed. In the next step the cover image is divided into the blocks of $c_h \times c_v$ by making use of the sequence of the numbers and the necessary route for the data embedding creation is done by block traversing from the right to the left and also from bottom to the top in the manner of the raster scan.

International Journal of Advance Research in Science and Engineering
Volume No.06, Issue No. 12, December 2017
www.ijarse.com
IJARSE
ISSN: 2319-8354

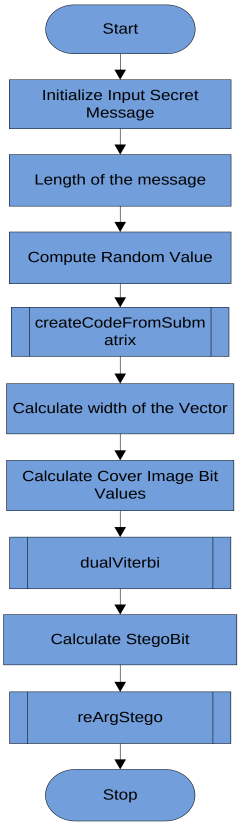


Fig.4. Flowchart for the Function Image Embedding

In the final step the data in the form of the binary is embedded into the pixel LSB along with the route created by the STC, also the seed selected is the one which produces distortion as the minimal and stored as the seed. The best seed for the generated PRNG which will result in the distortion which is minimal is determined. Finally for the creation of the stego image, the generation of the number sequence is again done by PRNG. With the help of the best seed the dividing of the cover image into the blocks are done, also the creation of the route for the embedding of the data is done and also the embedding of the binary data into the cover data by the help of the Syndrome Trellis code as in Figure 4. We should also make a note on the best seed for embedding into the region which is predefined of the stego image for the use of the future in the procedure of the extraction of the data.

2.4Data Extraction

The extraction of the data is done, which is must simpler than that of the embedding process. Every time the parameter c_v and c_h is shared in between the receiver and the sender. Initially the extraction of the best seed is done from the region which is predefined of the stego image obtained. In the next steps the obtained sequence $c_h \times (c_{v+1})$ number is generated by the PRNG which will be used in the procedure of the data embedding, with the seed which is best. In the next step the stego image division is done out of $c_h \times c_v$ blocks by making use of

International Journal of Advance Research in Science and Engineering
Volume No.06, Issue No. 12, December 2017
www.ijarse.com
IJARSE
ISSN: 2319-8354

the sequence of the numbers. At the end the creation of the route for the data extraction is done for the extraction of the binary data from the pixel LSB of the stego image along by the syndrome trellis codes.

III.EXPERIMENTAL RESULT

For the measurement of the performance in case of the image distortion, the PSNR i.e. signal to noise ratio will be applied for the stego image obtained. This given in the eq. (11),

$$PSNR = 10 \times \log_{10}\left(\frac{255^2}{MSE}\right) \qquad (11)$$

$$MSE = \frac{1}{hw}\sum_{x=1}^{h}\sum_{y=1}^{w}(S_{xy} - C_{xy})^2 \qquad (12)$$

Where the MSE i.e. mean square error is given by the eq. (12),

Where h denotes height and w denotes the width of the input image taken. S_{xy} Represents the location of the pixel of stego image and C_{xy} represents the location of the pixel value of the cover image.

Correlation: is a measure of similarity between the original cover image and stego image. The maximum value of correlation is 1. The original cover image and stego image are shown in figure 5, 6 observe both images are looking similar to each other even after embedding the secret bits into cover image.

Fig.5. (a) Input Cover Image

Fig.6. (a) Output Stego Image

Table1 shows the PSNR and MSE value for different image format bmp, jpg and png. The maximum PSNR value is obtained by using chilly image (bmp, png, jpg) is 70%. This shows that this algorithm is suitable for all type of images or it is independent of image formats. Measurement of entropy in steganography tells that the amount of deviation from pixel to pixel. By referring to the last column of the table 1 it is evident that the deviation of pixel is very less in the case of pepper.jpg it is zero. Correlation between stego and cover image is measured by and tabulated in column5 of the table1, shows that both images are not deviating much therefore avoiding the attention of intruders.

3.1 Histogram Analysis

Histogram is the graphical relation between pixel value and the frequency of occurrence. From figure 7 a) and b) it is observed that the histogram of cover image is same as the histogram of stego image, means there is not much variation in pixel value even after embedding the secret bits. Thus the quality of image is improved and it is suspicious from attacker or intruder.

Table1:For 512x512 image and 31936 secret bits.

Sol no	Cover image(512x512)	PSNR	MSE	Correlation	Entropy Difference
1	Baboon.bmp	69.7203	0.0069351	1.0000	0.0044
2	Baboon.png	69.7203	0.0069351	1.0000	0.0044
3	Baboon.jpg	69.8644	0.0067088	1.0000	0

4	Lena.bmp	67.9525	0.010419	1.0000	0.0751
5	Lena.png	67.9525	0.010419	1.0000	0.0751
6	Lena.jpg	69.9753	0.0065397	1.0000	0.0004
7	Chilly.bmp	70.3581	0.0059878	1.0000	0.0019
8	Chilly.png	70.3581	0.0059878	1.0000	0.0019
9	Chilly.jpg	70.181	0.006237	1.0000	0.0008
10	Pepper.bmp	69.7099	0.0069516	1.0000	0.0234
11	Pepper.png	69.7099	0.0069516	1.0000	0.0234
12	Pepper.jpg	69.9107	0.0066376	1.0000	0

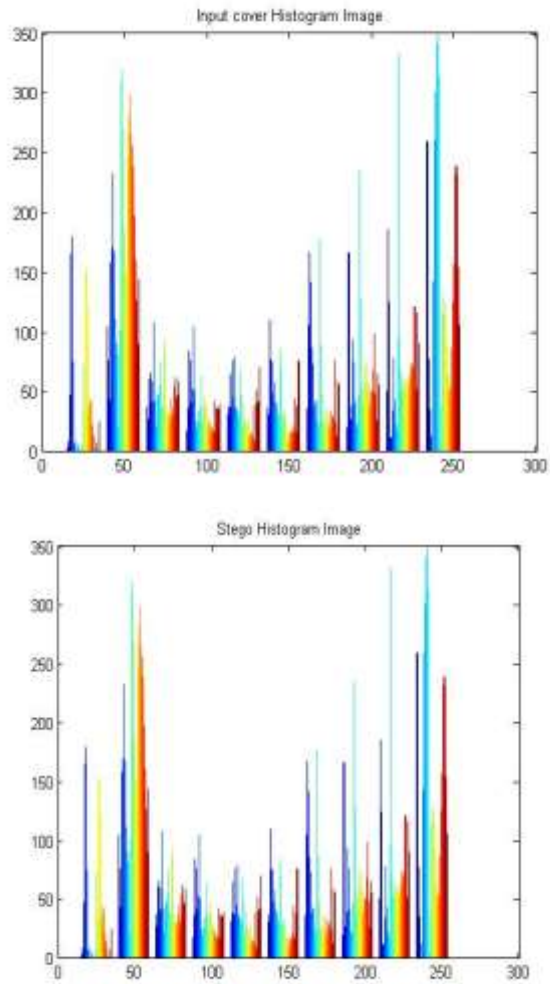


Fig.7. a) Histogram of input cover image b) Histogram of stego image

IV.CONCLUSION

Proposed methodology increases the embedding efficiency and minimizes the distortion by utilizing the STC (syndrome trellis code). We employed syndrome-trellis codes for sake of embedding data into the LSB of the pixels. The quality of stego image is analyzed by image histogram analysis. PRNG generates random number using a key Ch and Cv which is share at both receiver and sender side, without this key it is not possible to extract the secret bits .Thus Block generation algorithm increases the security performance.

REFERENCES

[1] Neha Rani, Jyoti Chaudhary International Journal of Engineering Trends and Technology (IJETT) – Volume 4 Issue 7- july 2013

[2] ArunaMalik,GeetaSikka, Harsh K. VermaEngineering Science and Technology,an International Journal 20 (2017) 72–79.

[3] Deepti Gupta, Swati Gupta et al, (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 2 (5) , 2011, 2060-2062 2060

[4] T. Filler, J. Judas and J. Fridrich, “Minimizing Additive Distortion in Steganography Using Syndrome-Trellis Codes,” IEEE, Vol. 6, Issue 3, pp. 920–935, 2011.

[5] Wei-Wei Liu, Guang-Jie Liu and Yue-Wei Dai, “Alternative Syndrome-Trellis Codes with Reduced Trellis Complexity Journal of Information Hiding and Multimedia Signal Processing”, Vol. 5, Issue 4, 2014.

[6] Iranpour, Mehran and Mohammad Rahmati, “A Novel Block-Based Steganographic Method”, IEEE, pp. 167-172, 2013.

[7] Spinner, Jens, Jürgen Freudenberger and SergoShavgulidze, “A soft input decoding algorithm for generalized concatenated codes”, IEEE, Vol. 64, Issue 9, pp. 3585-3595, 2016.

[8] Feng, Bingwen, Wei Lu and Wei Sun, “Secure Binary Image Steganography Based on Minimizing the Distortion on The Texture”, IEEE, Vol. 10, Issue 2, pp. 243-255, 2015.

[9] Guo, Linjie, Jiangqun Ni and Yun Qing Shi, “Uniform embedding for efficient JPEG steganography”, IEEE, Vol. 9, Issue 5, pp. 814-825, 2014.

[10] Xueyi, Y. E., Lu Guopeng, Wang Yunlu, and Zhang Yan, “A Jpeg Steganographic Method Based on Syndrome-Trellis Codes”, Journal of Theoretical & Applied Information Technology, Vol. 47, Issue 1, 2013.

[11] Filler, Tomáš, Jan Judas and Jessica Fridrich, “Minimizing embedding impact in steganography using trellis-coded quantization”, In IS&T/SPIE Electronic Imaging, pp. 754105-754105, 2010.

[12] ArunKumar.R, “Optimizing Distortion in Steganography Using S-T Codes”, International Journal of Advanced Research in Computer Science and Software Engineering”, Vol. 3, Issue 6, 2013.

[13] Malakooti, Mohammad V and MehrzadKhederzadeh, “ARobust and Lossless Information Embedding in Image Based on DCT and Scrambling Algorithms”, In TheInternational Conference on Computing Technologyand Information Management (ICCTIM), pp. 406-410, 2014.