HIGH SECURE DATA EMBEDDING USING STEGANOGRAPHY AND CRYPTOGRAPHY

P. Jeyalakshmi, R.Sankar

ABSTRACT

Cryptography and Steganography are the two major techniques for secret communication. The contents of secret message are scrambled in cryptography, where as in steganography the secret message is embedded into the cover medium. In this proposed system we developed high security model by combining cryptographic Steganographic security.This paper proposes a data hiding method based on pixel pair matching(PPM) with cryptographic technique. Here we applying a double transposition method with pixel pair matching.The basic idea of PPM is to use the values of pixel pair as a reference coordinate, and search a coordinate in the neighborhood set of this pixel pair according to a given message digit. The pixel pair is then replaced by the searched coordinate to conceal the digit.The proposed method always has lower distortion for various payloads and also secures under the detection of some well-known steganalysis techniques. Hiding data is better than moving it shown and encrypted.To hide data in a popular object that will not attract any attention In case the data is extracted; it will be encrypted.our proposed method enhances the security level.

***Keywords:** Cryptography,Steganography,Pixel pair matching(PPM),double transposition,steganalysis*

I. INTRODUCTION

DATA hiding is a technique that conceals data into a carrier for conveying secret messages confidentially.

Digital images are widely transmitted over the Internet; therefore, they often serve as a carrier for covert communication. Images used for carrying data are termed as cover images and images with data embedded are termed as stego images. After embedding, pixels of cover images will be modified and distortion occurs. The distortion caused by data embedding is called the embedding distortion . A good data-hiding method should be capable of evading visual and statistical detection while providing an adjustable payload [3]

The least significant bit substitution method, referred to as LSB in this paper, is a well-known data-hiding method. This method is easy to implement with low CPU cost, and has become one of the popular embedding techniques. However, in LSB embedding, the pixels with even values will be increased by one or kept unmodified. The pixels with odd values will be decreased by one or kept unmodified. Therefore, the imbalanced embedding distortion emerges and is vulnerable to steganalysis[6]. In 2004, Chan *et al.* [1] proposed a simple and efficient optimal pixel adjustment process (OPAP) method to reduce the distortion caused by LSB replacement. In their method, if message bits are embedded into the right-most LSBs of an r m-bit pixels other m-r bits are adjusted

[1]Research Scholar, M.E Computer Science Final Year, Kalasalingam Institute of Technology, Srivilliputtur.
India email : jeyalakshmi007@gmail.com,
[2]Research Scholar Ass. Prof., CSE Dept, Kalasalingam Institute Of Technology, Srivilliputtur. India
e-mail : sankarwt@gmail.com

by a simple evaluation.Namely, if the adjusted result offers a smaller distortion, these *m-r* bits are either replaced by the adjusted result or otherwise kept unmodified.

In 2006 Zhang and Wang [9] proposed an exploiting modification direction (EMD) method. EMD improves Mielikainen's method [6] in which only one pixel in a pixel pair is changed one gray-scale unit at most and a message digit in a 5-ary notational system can be embedded. Therefore, the payload is $(1/2)log_2 5=1.161$ bpp. LSB matching and EMD methods greatly improve the traditional LSB method in which a better stego image quality can be achieved under the same payload. However, the maximum payloads of LSB matching and EMD are only 1 and 1.161 bpp, respectively. Hence, these two methods are not suitable for applications requiring high payload. [4].

In 2009, Chao *et al.* [2] proposed a diamond encoding (DE) method to enhance the payload of EMD further. DE employs an extraction function to generate diamond characteristic values (DCV), and embedding is done by modifying the pixel pairs in the cover image according to their DCV's neighborhood set and the given message digit. Chao used an embedding parameter k to control the payload, in which a digit in a B -ary notational system can be concealed into two pixels,where $B=2K^2+2K+1$.*If* $k=1$,$B=5$,i.e.,digits in a 5-ary notational system are concealed,the resultant payload of EMD. Note that B is significantly increased as k is only increased by one. Instead of enhancing the payload of EMD, Wang *et al.* [7] in 2010 proposed a novel section-wise exploring modification direction method to enhance the image quality of EMD. Their method segments the cover image into pixel sections, and each section is partitioned into the selective and descriptive groups [5].

PPM based data embedding method reduce the embedding impact by providing a simple extraction function and a more compact neighborhood set. This method embeds more messages per modification and thus increases the embedding efficiency. The image quality obtained by this method not only performs better than those obtained by OPAP and DE, but also brings higher payload with less detectability. Moreover, the best notational system for data concealing can be determined and employed in this method according to the given payload so that a lower image distortion can be achieved.

II. Adaptive pixel pair matching (APPM)

The basic idea of the PPM-based data-hiding method is to use pixel pair (x,y) as the coordinate, and searching a coordinate (x',y') within a predefined neighborhood set $\varphi(x,y)$ such that $f(x',y')=S_B$, where f is the extraction function and S_B is the message digit in a B -ary notational system to be concealed. Data embedding is done by replacing (x,y) with (x',y').

For a PPM-based method, suppose a digit S_B is to be concealed. The range of S_B is between 0 and B-1 , and a coordinate (x',y') $\varphi(x,y)$ has to be found such that $f(x',y')=S_B$.Therefore, the range *of* $f(x,y)$ must be integers between 0 and B-1 , and each integer must occur at least once. In addition, to reduce the distortion, the number of coordinates in φ(x,y) should be as small as possible. The best PPM method shall satisfy the following three requirements: 1) There are exactly B coordinates in φ(x,y). 2) The values of extraction function in these coordinates are mutually exclusive. 3) The design of φ(x,y) and f(x,y) should be capable of embedding digits in any notational system so that the best B can be selected to achieve lower embedding distortion[8]

DE is a data-hiding method based on PPM. DE greatly enhances the payload of EMD while preserving acceptable stego image quality. However, there are several problems. First, the payload of DE is determined by the selected notational system, which is restricted by the parameter K; therefore, the notational system cannot be arbitrarily selected. For example, when k is 1, 2, and 3, then digits in a 5-ary, 13-ary, and 25-ary notational system are used to embed data, respectively. However, embedding digits in a 4-ary (i.e., 1 bit per pixel) or 16-ary (i.e., 2 bits per pixel) notational system are not supported in DE. Second ,in DE is defined by a diamond shape, which may lead to some unnecessary distortion when K>2.. Infact, there $\varphi(x,y)$ exists a better $\varphi(X,Y)$ other than diamond shape resulting in a smaller embedding distortion.This method not only allows concealing digits in any notational system, but also provides the same or even smaller embedding distortion than DE for various payloads.

A. Extraction function and neighthood set

The definitions of $\varphi(x,y)$and $f(x,y)$ significantly affect the stego image quality. The designs of φ(x,y) and $f(x,y)$ have to fulfill the requirements: all values of $f(x,y)$ in φ(x,y) have to be mutually exclusive, and the summation of the squared distances between all coordinates in $\varphi(x,y)$ and (x,y) has to be the smallest. This is because, during embedding, (x,y) is replaced by one of the coordinates in φ(x,y) Suppose there are B coordinates in φ(x,y),i.e., digits in a B-ary notational system are to be concealed, and the probability of replacing (x,y) by one of the coordinates in φ(x,y) is equivalent.

The averaged MSE can be obtained by averaging the summation of the squared between (x,y) and other coordinates in φ(x,y) .Thus,given a φ(x,y), the expected MSE after embedding can be calculated by

$$\mathrm{MSE}{=}1/2\mathrm{B}\sum_{i=0}^{B-1}(x_i-x)^2+(y_i-y)^2$$

This adaptive pixel pair matching(APPM) data hiding method to explore better $f(x,y)$ and $\varphi_{..}$(x,y) so that MSE$\varphi_{(x,y)}$ is minimized.data is then embedded by using PPM based on these f(x,y) and φ (x,y).Let

$$f(x,y)=(x+c_B * y) \bmod B$$

The solution of ,..(x,y) and $f(x,y)$ is indeed a discrete optimization problem

$$\text{Minimize:} \quad \sum_{i=0}^{B-1}(x_i-x)^2+(y_i-y)^2$$

$$\text{Subject to:} \quad f(x_i,y_i) \quad \{0,1,\ldots\mathrm{B}\text{-}1\} \qquad (1)$$

$$f(x,y) \neq f(x,y), \ \text{if } i \neq j$$

$$\text{for } 0 \leq i,, j \leq \mathrm{B}\text{-}1.$$

Given an integer B and an integer pair (x,y) , (1) can be solved to obtain a constant C_B and B pairs of (x_i,y_i).These B pairs of $(x_i,y_i)$are denoted by φ_B(x,y). Note that φ_B (x,y) represents a neighborhood set of (x,y) . Table I lists the constant C_B satisfying (1) for the payloads under 3 bpp. Note that, for a given B, it is possible to have more than one C_B and φ_B (x,y) , satisfying (1). Table I only lists the smallest C_B.

Fig.1 shows some representative φ_B(x,y) and their corresponding C_B satisfying (1), where the center of φ_B(x,y) is shaded with lines. Note that, in DE, setting

Table:1 C_B Values for 2d≤Rd≤64

C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_{10}	C_{11}	C_{12}	C_{13}	C_{14}	C_{15}	C_{16}	C_{17}	C_{1}
1	1	2	2	2	2	3	3	3	3	4	5	4	4	6	4	4
C_{19}	C_{20}	C_{21}	C_{22}	C_{23}	C_{24}	C_{25}	C_{26}	C_{27}	C_{28}	C_{29}	C_{30}	C_{31}	C_{32}	C_{33}	C_{34}	C_{3}
4	8	4	5	5	5	5	10	5	5	5	12	12	7	6	6	10
C_{36}	C_{37}	C_{38}	C_{39}	C_{40}	C_{41}	C_{42}	C_{43}	C_{44}	C_{45}	C_{46}	C_{47}	C_{48}	C_{49}	C_{50}	C_{51}	C_{5}
15	6	16	7	7	6	12	12	8	7	7	7	7	14	14	9	22
C_{53}	C_{54}	C_{55}	C_{56}	C_{57}	C_{58}	C_{59}	C_{60}	C_{61}	C_{62}	C_{63}	C_{64}					
8	12	21	16	24	22	9	8	8	8	14	14					

k=3 and k=4,respectively, embeds digits in the 25-ary and 41-ary notational systems. We also depict the φ(x,y) of DE when setting k=3, and k=4. in Fig. . Note that the four corners of the diamond shape may cause larger MSE but ours selects a more compact region for embedding, and thus smaller distortion can be achieved.

	3	4		
	9	10	11	
14	15	0	1	2
	5	6	7	8
		12	13	

Figure 1 φ_{16} (0,0) Neighbourhood set

x_0,y_0)=(0,0)
(x_1,y_1)=(1,0)
(x_2,y_2)=(2,0)
.
.
.
(x_5,y_5)=(-1,1)
.
(x_{15},y_{15})=(-1,0)

III. Double transposition cipher

To encrypt with Double Transposition Cipher, we first write the plaintext into an array of a given size and then permute rows and columns according to specified permutations[10]. for example suppose we write a plaintext attackatdawn into 3×4 array

$$\begin{bmatrix} a & t & t & a \\ c & k & a & t \\ d & a & w & n \end{bmatrix}$$

Now we transpose (or) permute the rows according to (1 2 3) ⟶ (3 2 1) and then transpose the columns according to (1 2 3 4) ⟶ (4 2 1 3),we obtain

$$\begin{bmatrix} a & t & t & a \\ c & k & a & t \\ d & a & w & n \end{bmatrix} \longrightarrow \begin{bmatrix} d & a & w & n \\ c & k & a & t \\ a & t & t & a \end{bmatrix}$$

$$\begin{bmatrix} n & a & d & w \\ t & k & c & a \\ a & t & a & t \end{bmatrix} \longleftarrow$$

Ciphertext is then read from final array "nadwtkcaatat".

For double transposition,the key consists of the size of the matrix and row and column permutations.the recipient who knows the key can simply put the

$$\begin{tikzcd} & \begin{bmatrix} n & a & d & w \\ t & k & c & a \\ a & t & a & t \end{bmatrix} \arrow[r] & \begin{bmatrix} d & a & w & n \\ c & k & a & t \\ a & t & t & a \end{bmatrix} \arrow[dll, bend left] \\ \begin{bmatrix} a & t & t & a \\ c & k & a & t \\ d & a & w & n \end{bmatrix} & & \end{tikzcd}$$

ciphertext into the appropriate sized matrix undo the the permutations to recover the plaintext,for example

to decrypt ciphertext,the ciphertext first put into the 3×4 array then the columns are numbered as (4 2 1 3) and rearranged to (1 2 3 4) then the rows are numbered (3 2 1) and rearranged into (1 2 3),as illustrated below.

And we have recovered plain text,"attackatdawn".the double transposition is not a trivial cipher,to break the idea of "smearing" plaintext information through the ciphertext is so useful that it is employed by modern block ciphers.Here we using elliptic curve cryptography.it provides confidentiality and security.

IV. Modules

1) Encoding procedure

2) Embedding procedure

3) Extraction procedure

4) Decoding procedure

A. Encoding procedure

In Encoding procedure ,we convert a plain text into cipher text using Double transposition technique,then we embedding cipher text into cover image.

B. Embedding procedure

Suppose the cover image is of size M × M ,S is the message bits to be concealed and the size of S is |S|. First we calculate the minimum B such that all the message bits can be embedded. Then, message digits are sequentially concealed into pairs of pixels. The detailed procedure is listed as follows.

Input : Cover image I of size M × M,secret bit stream S,and key K_r

Output:Stego image I', φ_B (x,y) , C_B, and K_r.

1. Find the minimum Bsatisfying [M × M/2] ≥ $|S_B|$ and convert S into a list of digits with a , B -ary notational system S_B
2. Solve the discrete optimization problem to find C_B and , φ_B (x,y).
3. In the region defined by φ_B (0,0) , record the coordinate (x_i,y_i) such that $f(x_i,y_i)$=i, 0≤i≤B-1.
4. Construct a nonrepeat random embedding sequence Q using a key K_r
5. To embed a message digit S_B , two pixels *(x,y)* in the cover images are selected according to the embedding sequence Q,and calculate the modulus distance d = (S_B – f(x,y)) mod B between S_B and *f(x,y),* then replace *(x,y)* with $(x+x_d,y+y_d)$

6.. Repeat Step 5 until all the message digits are embedded.

In real applications, we can solve all C_B and φ_B (x,y) at once. With the knowledge of C_B and φ_B (x,y), there is no need to perform Step 2 in the embedding phase.

Let $x'=x+x_d$ and $y'=y+y_d$..if an overûow or underûow problem occurs, that is, (x',y') < 0or (x',y') > 255,then in the neighborhood of (x,y)find a nearest (x",y") such that f(x",y")=S_B This can be done by solving the optimization problem

Minimize

$$(x-x'')^2+(y-y'')^2$$

Subject to $f(x'',y'')=S_B$,0≤ x",y"≤ 255

C. Extraction procedure

To extract the embedded message digits, pixel pairs are scanned in the same order as in the embedding procedure. The embedded message digits are the values of extraction function of the scanned pixel pairs.

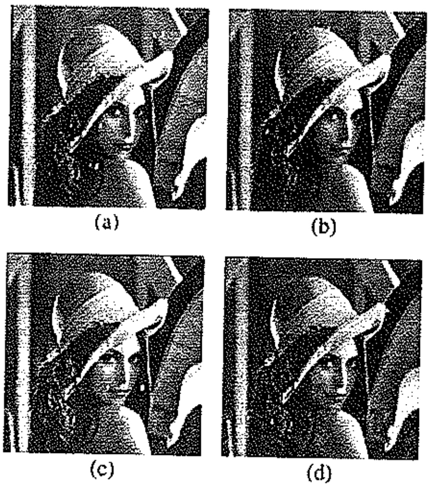
Input: Stego image I', φ_B (x,y) , $C_{B,}$ and $K_{r.}$

Output: Secret bit stream S

1. Construct the embedding sequence Q using the key K_r.
2. Select two pixels (x',y')according to the embedding sequence Q
3. Calculate,*f(x',y')* the result is the embedded digit
4. Repeat Steps 2 and 3 until all the message digits are extracted.
5. Finally, the message bits S can be obtained by converting the extracted message digits into a binary bit stream.

D. Decoding procedure

In the Decoding procedure,Cipher text is converted into plain text.That means After extracting a cipher text from stego image,that text was converted into original plain text.



V. QUALITY ANALYSIS FOR APPM

Figure 2. Cover image and stego images under various payloads. (a) Cover image. (b) Stego image, 2 bpp at 46.86 dB. (c) Stego image, 3 bpp at 40.97 dB. (d) Stego image, 4 bpp at 34.90 dB.

I) Mse comparison of APPM

Figure 2 explains about cover image and stego image under various payloads.MSE comparison of some PPM-based data- hiding methods for payload less than 2 bpp. It can be seen that the MSEs of APPM are always smaller or equal to other PPM- based methods. For example, when digits in a 4-ary notational system are embedded, the MSEs of APPM and LSB matching are the same. When embedding digits in a 13-ary notational system, APPM and DE (k=2) have the same MSE. However, when embedding 16-ary digits, APPM outperforms OPAP. APPM not only greatly increases the payload of EMD, but also enable users to freely select the desired notational system for data embedding so that a better image quality can be obtained.

A. Comparison with other methods

The performance of the APPM method is the best under various payloads. For example with the payload 400 000 bits, the averaged MSE of 2-bit OPAP is 1.244, whereas the averaged MSE of DE is 0.887. However, this APPM method has the smallest averaged MSE, 0.640. For larger payload, such as 650 000 and 1 000 000 bits, this method also performs better than OPAP and DE because APPM selects the smallest notational system that provides just enough embedding capacity to accommodate the given payload with the least distortion.

VI. Conclusion

This paper proposed a double transposition cryptographic technique with simple and efficient data embedding method based on PPM. Two pixels are scanned as an embedding unit and a specially designed neighborhood set is employed to embed message digits with a smallest notational system. APPM allows users to select digits in any notational system for data embedding, and thus achieves a better image quality.

This method not only resolves the low-payload problem in EMD, but also offers smaller MSE compared with OPAP and DE. Moreover, because APPM produces no artifacts in stego images and the steganalysis results are similar to those of the cover images, it offers a secure communication under adjustable embedding capacity. We combine cryptographic technique with this method.so it enhances the security level.

References

[1] Chan.C. K. and Cheng .L. M., *"Hiding data in images by simple LSB substitution,"* Pattern Recognit., vol. 37, no. 3, pp. 469–474, 2004.

[2] Chao R.M., Wu H. C., Lee C. C., and Chu Y. P., *"A novel image data hiding scheme with diamond encoding," EURASIP J. Inf. Security*, vol.2009, 2009, DOI: 10.1155/2009/658047, Article ID 658047.

[3] Fridrich .J. and Filler .T., *"Practical methods for minimizing embedding impact in steganography,"* in Proc. SPIE, Security, Stegano graphy., Water marking of Multimedia, 2007, vol. 6050, pp.2–3.

[4] Hong .W and Chen T. S., *"Reversible data embedding for high quality images using interpolation and reference pixel distribution mechanism,"*J. Vis. Commun. Image Represent., vol. 22, no. 2, pp. 131–140, 2011

[5] Ker. A. D., *"Steganalysis of LSB matching in grayscale images,"* IEEE Signal Process. Lett., vol. 12, no. 6, pp. 441–444, Jun. 2005.

[6] Mielikainen.J, *"LSB matching revisited,"* IEEE Signal Process. Lett.,vol. 13, no. 5, pp. 285–287, May 2006.

[7] Wang.R, Sun. Y, Xu .H.,. Chen.K, H. J. Kim, and S. H. Joo, *"An improved section-wise exploiting modification direction method,"* Signal Process., vol. 90, no. 11, pp. 2954–2964, 2010.

[8] Zhao .H., Wang .H., and Khan M. K., *"Statistical analysis of several reversible data hiding algorithms,"* in Proc. Multimedia Tools and Applications,2009, DOI: 10.1007/s11042-009-0380-y..

[9] Zhang.X. and Wang .S., *"Efficient steganographic embedding by exploiting modification direction,"* IEEE Commun. Lett., vol. 10, no. 11,pp. 781–783, Nov. 2006.

[10] William Stallings," Cryptography and Network Security"Fifth Edition.

Authors Biography

P. Jeyalakshmi, completed her school education in S.R.N Hr. Sec School, Thiruthangal. She did her Bachelor Of Engineering in Park College Of Engineering and technologyin kovai. She is presently doing her final year Master Of Engineering in Kalasalingam Institute Of Technology, Virudhunager. Her area of interest is network Security and has presented her paper in three National Level Conference and one international conference.