AN EFFICIENT DATA EMBEDDING TECHNIQUE USING IMAGE AS A DIGITAL MEDIA

Pavithra S R*

V.Venkateswara Reddy**

ABSTRACT

In the present world of communication, one of the necessary requirements to prevent data theft is securing the information. Security has become a critical feature for thriving networks and in military alike. Cryptography and Steganography are well known and widely used techniques that manipulate information (message) in order to cipher or hide their existence. These techniques have many applications in computer science and other related fields: they are used to protect military messages, E-mails, credit card information, corporate data, personal files, etc. Data hiding is a technique that conceals data into a carrier signal for conveying secret messages confidentially. Steganography hides the message in digital media. The purpose of steganography is to conceal the fact that some communication is taking place. 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 The basic idea of pixel pair matching 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. Exploiting modification direction (EMD) and diamond encoding (DE) are two data-hiding methods proposed recently based on PPM. In the EMD method 5-ary method is used as notational system so its payload is 1.161 bpp and DE extends the payload of EMD by embedding digits in a larger notational system. The APPM method offers lower distortion than DE by providing more compact neighborhood sets and allowing embedded digits in any notational system. Image distortion occurs when pixel values are modified because of data embedding. We use MSE to measure the image quality. MSE represents the mean square error between the cover image and stego image. A smaller MSE indicates that the stego image has better image quality.

***Keywords**: PPM (pixel pair matching), MSE (mean square error), cover image, stego image*

*IV Sem, M.Tech (Signal processing), Dept. of E&C, SJCIT, Chickballapur, Karnataka, India

**Assistant Professor, Dept. of E&C, SJCIT, Chickballapur, Karnataka, India

I. INTRODUCTION

In the present world of communication, one of the necessary requirements to prevent data theft is securing the information [1]. Security has become a critical feature for thriving networks and in military alike. Cryptography and Steganography are well known and widely used techniques that manipulate information (message) in order to cipher or hide their existence. These techniques have many applications in computer science and other related fields: they are used to protect military messages, E-mails, credit card information, corporate data, personal files, etc.

Steganography hides the message in innocent digital file [1]. The purpose of steganography is to conceal the fact that some communication is taking place [3] [2]. With any type of hidden communication, the security of the message often lies in the secrecy of its existence and/or the secrecy of how to decode it. The steganography hides the secret information behind a cover so that it draws no special attention [6]. The cover represents any digital file like image, text, and video, sound and …, etc. If we used the digital image, the cover-image after embedding is called stego-image [1].

Data hiding is a technique that conceals data into a carrier for conveying secret messages confidentially [1], [2]. 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 [3]. A good data-hiding method should be capable of evading visual and statistical detection [4] while providing an adjustable payload [5].

Steganography is applicable to

1. Confidential communication and secret data storing,
2. Protection of data alteration,
3. Access control system for digital content distribution,

Steganalysis is the science of detecting hidden information [1]. The main objective of Steganalysis is to break steganography and the detection of stego image is the goal of Steganalysis. Almost all Steganalysis algorithms rely on the Steganographic algorithms introducing statistical differences between cover and stego image. Steganalysis deals with three important categories [1, 4].

(a) Visual attacks: In these types of attacks with a assistance of a computer or through inspection with a naked eye it reveal the presence of hidden information, which helps to separate the image into bit planes for further more analysis.

(b) Statistical attacks: These types of attacks are more powerful and successful, because they reveal the smallest alterations in an images statistical behavior.

Statistical attacks can be further divided into (i) Passive attack and (ii) Active attack.

Passive attacks involves with identifying presence or absence of a covert message or embedding algorithm used etc.

Mean while active attacks is used to investigate embedded message length or hidden message location or secret key used in embedding.

(c) Structural attacks: The format of the data files changes as the data to be hidden is embedded; identifying this characteristic structure changes can help us to find the presence of image.

II. LITERATURE SURVEY:

In 2004, Chan et al. [6] proposed a simple and efficient optimal pixel adjustment process (OPAP) method to reduce the distortion caused by LSB replacement. In that method the pixels with even values will be increased by one. The pixels with odd values will be decreased by one. If the adjusted result offers a smaller distortion, LSB bits are either replaced by the adjusted result or otherwise kept unmodified. In 2006, Mielikainen [7] proposed an LSB matching method based on PPM. He used two pixels as an embedding unit. The LSB of the first pixel is used for carrying one message bit, while a binary function is employed to carry another bit. In the same year, Zhang and Wang [8] proposed an exploiting modification direction (EMD) method. EMD improves Mielikainen's method 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) $\log_2 5$=1.16 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 .In 2008, Hong [9] presented a data-hiding method based on Sudoku solutions to achieve a maximum payload of (1/2)$\log_2 9$ bpp. In 2009, Chao et al. [10] 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. Wang et al. [11] 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. The EMD embedding procedure is then performed on each group by referencing a predefined selector and descriptor table. This method combines different pixel groups of the cover image to represent more embedding directions with less pixel changes than that of the EMD method. By selecting the appropriate combination of pixel groups, the embedding efficiency and the visual quality of the stego image is enhanced.

In this project proposes a new data embedding method to reduce the embedding impact by providing a simple extraction function and a more compact neighborhood set. The proposed method embeds more messages per modification and thus increases the embedding efficiency. The image quality obtained by the proposed method not only performs better than those obtained by OPAP and DE, but also brings higher payload with less detectability.

III. RELATED WORK:

A. LSB substitution method:

1. The pixels with even values will be increased by one or kept unmodified.

Ex: If the pixel value is **36 →00110110**

After applying LSB substitution method for data embedding

Pixel value will be equal to **37→00110111(**increased by one**)**

2. The pixels with odd values will be decreased by one or kept unmodified.

Ex: If the pixel value is **35→00110101**

After applying LSB substitution method for data embedding

Pixel value will be equal to **34→00110100(**decreased by one**)**

If we take another example:

A pixel (225,100,100) with character "a", then we can obtain:

Original pixel = (11100001, 01100100, 01100100) "a" = 01100001(ASCII value 97)

New pixel = (11100011, 01100000, 01100101)

New pixel = (227, 96,101),

Here we can notice that the new pixel of (227, 96,101) is almost the same value as the old pixel of (225,100,100). So there will not be noticeable color difference in the image.

B. OPAP (Optimal Pixel Adjustment Method):

Optimal pixel adjustment process (OPAP) is proposed to enhance the image quality of the stego-image obtained by the simple LSB substitution method.

Let $\boldsymbol{p_i}$, $\boldsymbol{p_i'}$ and $\boldsymbol{p_i''}$ be the corresponding pixel values of the i^{th}pixel in the cover-image *C*, the stego-image C' obtained by the simple LSB substitution method and the refined stego-image obtained after the OPAP. Let $\boldsymbol{\delta_i = p_i' - p_i}$ be the embedding error between p_i and p_i'. According to the embedding process of the simple LSB substitution method is obtained by the direct replacement of the *k* least significant bits of p, with *k* message bits, therefore, $\boldsymbol{-2^k < \delta_i < 2^k}$.

The value of δ_i, can be further segmented into three intervals. Such that,

Interval 1: $-2^{k-1} < \delta_i < 2^k$,

Interval 2: $-2^{k-1} \le \delta_i \le 2^{k-1}$,

Interval 3: $-2^k < \delta_i < -2^{k-1}$.

Based on the three intervals, the OPAP, which modifies p_i' to form the stego-pixel p_i'' can be described as follows:

Case I

$(-2^{k-1} < \delta_i < 2^k)$:

If $p_i' \ge 2^k$, then $p_i'' = p_i' - 2^k$;

Otherwise $p_i'' = p_i'$;

Case II

$(-2^{k-1} \le \delta_i \le 2^{k-1})$:

$p_i'' = p_i'$;

Case III

$(-2^k < \delta_i < -2^{k-1})$:

If $p_i' < 256{-}2^k$ then $p_i'' = p_i' + 2^k$;

Otherwise $p_i'' = p_i' - 2^k$.

C. EMD (Exploiting Modification Direction):

In which only one pixel pair is changed one grayscale unit at most and a message digit in a 5-ary notational system can be embedded.

That is pixel values are represented in 5-ary notational system before embedding the data in that pixel.

Ex: Take an pixel with value (1101 0110 1001)

This is represented in 5-ary as **(23 11 14)**

For 1101 → 13 → 13+10=23(if the resulting decimal value is one digit ,we should add 5(5*1) to that value ,similarly if the if the resulting decimal value is two digit ,we should add 10(5*2) to that value and so on)

For 0110 → 6 → 6+5=11

For 1001 → 9 → 9+5=14

EMD method offers the fixed payload, because in this method notational system is fixed for 5-ary.

Therefore,

Payload= $(1/2)\log_2 5$=1.161 bpp (bits per pixel)

D. Diamond Encoding Method (DE)

This method is purely based on pixel pair matching. This method conceals the data in a B-ary system in to two pixels.

Where,

B =$2k^2 + 2k + 1$, $k \geq 1$..........................(1)

Therefore,

Payload= $(1/2)\log_2(2k^2 + 2k + 1)$ bpp......(2)

Where,

k → is embedding parameter to control payload.Then, DE method greatly enhances the payload of EMD while preserving an acceptable stego image quality.

For k=1,

DE=EMD (because for k =1, notational system is 5-ary)

Using equation (2),

Payload = $(1/2)\log_2(2k^2 + 2k + 1)$

=$(1/2)\log_2(2 + 2 + 1)$ =1.161 bpp

	p,q-1 **(0,-1)**	
p-1,q **(-1,0)**	**p,q** **(0,0)**	**p+1,q** **(1,0)**
	p,q+1 **(0,1)**	

For k=2,

Payload = $(1/2)\log_2(2k^2 + 2k + 1)$

= $(1/2)\log_2(2(4) + 2(2) + 1)$

=1.850 bpp

		p,q-2		
	p-1,q-1	p,q-1	p+1,q-1	
p-2,q	p-1,q	p,q	p+1,1	p+2,q
	p-1,q+1	p,q+1	p+1,q+1	
		p,q+2		

The neighborhood set (x, y) is represented as,

$$(x, y) = \{(x, y)|\ |a - x| + |b - y| \leq \qquad k\} \ldots\text{..(3)}$$

Where,

(x, y)represents the set of the coordinates(a, b)'s whose absolute distance to the coordinate (x, y) is smaller or equal to k.

A diamond function 'f' is then employed to calculate the DCV (diamond characteristic values) of (x, y),

Therefore,

$$f(x, y) = \big((2k + 1)x + y\big)\ mod\ B \ldots\text{..(4)}$$

Where,

Payload = $(1/2)\log_2 13$=1.85 bpp

To calculate the **DCV** (diamond characteristic values) of (x, y) using equation (4),

$$f(0,0) = \big((5) * 0 + 0\big)mod\ (13) = \mathbf{0}$$

$$f(0, 1) = \big((5) * 0 + 1\big)mod\ (13) = \mathbf{1}$$

$$f(0, 2) = \big((5) * 0 + 2\big)mod\ (13) = \mathbf{2}$$

$$f(0, -1) = \big((5) * 0 - 1\big)mod\ (13) = \mathbf{12}$$

$$f(0, -2) = \big((5) * 0 - 2\big)mod\ (13) = \mathbf{11}$$

$$f(1, 0) = \big((5) * 1 + 0\big)mod\ (13) = \mathbf{5}$$

$$f(1, 1) = \big((5) * 1 + 1\big)mod\ (13) = \mathbf{6}$$

$$f(1, -1) = \big((5) * 1 - 1\big)mod\ (13) = \mathbf{4}$$

$$f(-1, 0) = \big((5) * (-1) + 0\big)mod\ (13) = \mathbf{8}$$

$$f(-1, 1) = \big((5) * (-1) + 1\big)mod\ (13) = \mathbf{9}$$

$$f(-1, -1) = \big((5) * (-1) - 1\big)mod\ (13) = \mathbf{7}$$

$$f(2, 0) = \big((5) * 2 + 0\big)mod\ (13) = \mathbf{10}$$

$$f(-2, 0) = \big((5) * (-2) + 0\big)mod\ (13) = \mathbf{3}$$

This can be shown in a diamond structure as,

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

Ex: For k=2

We know that,

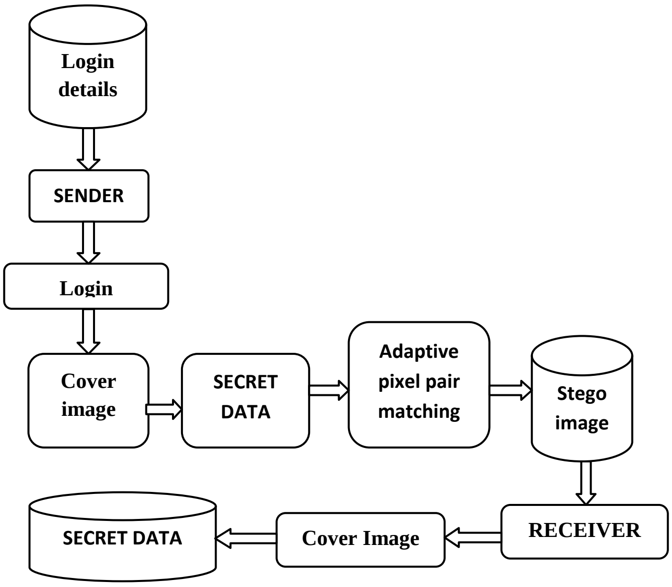
$B=2k^2+2k+1\,,\ \ k \geq 1$(from equation (1))

Therefore, B=2(4) + 2(2) + 1 = **13.**

Disadvantages of DE method:

However embedding digits in a 4-ary (i.e, 1-bit per pixel) or 16-ary(i.e, 2-bits per pixel) are not supported in DE method, because, the value of notational system is fixed in diamond encoding method and it depends completely on the K's value.

IV. SYSTEM ARCHITECTURE



A. Existing System

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. Optimal pixel adjustment process (OPAP) method to reduce the distortion caused by LSB replacement. In their method if message bits are embedded into the rightmost LSBs of an 2-bit pixel, other bits are adjusted by a simple evaluation. Namely, if the adjusted result offers a smaller distortion, these bits are either replaced by the adjusted result or otherwise kept unmodified. Exploiting modification direction (EMD) and diamond encoding (DE) are two data-hiding methods proposed recently based on PPM.

Disadvantages of existing system

- Imbalanced embedding distortion emerges and is vulnerable to steganalysis.
- The existing technique can be easily cracked.

B. Proposed System

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

Advantages of Proposed System:

The proposed method offers lower distortion than DE by providing more compact neighborhood sets and allowing embedded digits in any notational system. Compared with the optimal pixel adjustment process (OPAP) method, the proposed method always has lower distortion for various payloads. Experimental results reveal that the proposed method not only provides better performance than those of OPAP and DE. But also is secure under the detection of some well known steganalysis techniques.

V. MODULES

1) Extraction Function and Neighborhood Set.
2) Embedding Procedure.
3) Extraction Procedure

4) Statistical Analysis

Modules Description

1) Extraction function and neighborhood set

In this module we perform the action of extraction function and neighborhood set. Where the system does a new data embedding method to reduce the embedding impact by providing a simple extraction function and a more compact neighborhood set. The proposed method embeds more messages per modification and thus increases the embedding efficiency. The image quality obtained by the proposed 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 new method according to the given payload is that a lower image distortion can be achieved.

2) Embedding Procedure

Input: Cover image of size, secret bit strewn, and key.

Output: Stego image. and.

1. Find the minimum satisfying, and convert into a list of digits with a •ary notational system.
2. Solve the discrete optimization problem to find and.
3. In the region defined by record the coordinate such that..
4. Construct a no repeat random embedding sequence using a key.
5. To embed a message digit. two pixels in the cover image are selected according to the embedding sequence. and calculate the modulus distance between and then replace with.
6. Repeat Step 5 until all the message digits are embedded.

3) 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., . and.

Output: Secret bit stream.

1. Construct the embedding sequence using the key.
2. Select two pixels according to the embedding sequence.
3. Calculate, the result is the embedded digit.
4. Repeat Steps 2 and 3 until all the message digits are extracted.

5. Finally, the message hits can be obtained by converting the extracted message digits into a binary hit strewn.

4) Statistical Analysis of the Histogram Differences

In this module, we perform the goal of system analysis by using histogram technique. The goal of steganography is to evade statistical detection. It is apparent that MSE is not a good measure of security against the detection of steganalysis.

$$\boldsymbol{MSE}_{\Phi(\mathrm{x,y})} = \frac{\mathbf{1}}{\mathbf{2B}}\sum_{i=0}^{B-1}((x_i - x)^2 + (y_i - y)^2) \ldots.(5)$$

Histograms are used to plot density of data, and often for density estimation: estimating the probability density function of the underlying variable. The total area of a histogram used for probability density is always normalized to 1. If the lengths of the intervals on the x-axis are all I. then a histogram is identical to a relative frequency plot.

MSE comparison of the proposed method with LSB and OPAP

Payload (bpp)	LSB	OPAP	APPM	MSE improvement over OPAP
1	0.500	0.500	.375	0.125
2	2.500	1.500	1.344	0.156
3	10.50	5.500	5.203	0.297
4	42.50	21.50	20.51	0.982

VI. CONCLUSION

- Simple and efficient
- APPM allows users to select digits in any notation system embedding
- Offer small MSE compared with OPAP and DE

VII. FUTURE WORK

It is expected that our adaptive idea can be extended to other steganographic methods such as audio/video steganography in the spatial or frequency domains

VII. REFERENCES

[1] Ingemar J. Cox, J.Fridrich, “Digital Watermarking and Steganography” second edition, Morgan Kaufmann Publishers, Press, 2008.

[2] N. Provos and P. Honeyman, “Hide and seek: An introduction to steganography,” *IEEE Security Privacy*, vol. 3, no. 3, pp. 32–44, May/Jun. 2003.

[3] A. Cheddad, J. Condell, K. Curran, and P. McK evitt, “Digital image steganography: Survey and analysis of current methods,” *Signal Process.*, vol. 90, pp. 727–752, 2010.

[4]Saurabh Singh,Gaurav Agarwal“Use of image to secure text message with the help of L SB replacement,”International journal of applied engineering research, dindigul volum e 1, no1, 2010

[5] S. Lyu and H. Farid, “Steganalysis using higher-order image statistics,”*IEEE Trans. Inf. Forensics Security*, vol.1, no. 1, pp. 111–119, Mar.2011.

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

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

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

[9] W. Hong, T. S. Chen, and C. W. Shiu, “A minimal Euclidean distance searching technique for Sudoku steganography,” in *Proc. Int. Symp. Information Science and Engineering*, 2008, vol. 1, pp. 515–518.

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

[11] J. Wang, Y. Sun, H. Xu, K. Chen, 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.