Data Embedding Techniques in Steganography

K.Saranya

II M.E, Computer science and Engineering, Vivekanandha College of Engineering for Women, Tamilnadu, India

Dr.C.Suresh Gnanadhas

Professor, Computer science and Engineering, Vivekanandha College of Engineering for Women, Tamilnadu, India

Minu George

II M.E, Computer science and Engineering, Vivekanandha College of Engineering for Women, Tamilnadu, India

Abstract - Steganography is the technology of hiding data into digital media without drawing any suspiction. It can be used to carry out hidden exchanges of information and to provide privacy for the user The messages are inserted into the cover image by using stego – key. The main issues in Steganography are capacity and security. A secure and efficient mechanism must be provided before covering the information. Data embedding technique is one of the method used for hide the messages over the image. In this paper, we are going to analysis various algorithms in data embedding technique that are Edge Adaptive image Steganography based on LSB Matching Revisited, High – dimensional image models to perform highly undetectable Steganography, the amplitude of histogram local extrema, LSB Matching in image with high – frequency noise. These algorithm aims to achieve more security and to improve the stego image quality.

Keywords: Information hiding, Stego image, Attacks, Pseudorandom Number Generator (PNG)

I. INTRODUCTION

Data hiding is a technique that conceals data into a carrier for conveying secret messages confidentially. The medium in which the message is embedded is referred as cover image and the resulting image and the message combined is referred as stego–image. Pixels of cover images will be modified after data embedding and also distortion occurs. The notion of distortion caused by data embedding is called the embedding distortion. In LSB embedding, the pixels with even values will be increased by one or more even it kept unmodified. The pixels with odd values will be decreased by one or more even it kept unmodified. LSB replacement is a well – known steganographic technique. In this embedded scheme, only LSB plane of the cover image is overwritten with the secret bit stream according to a pseudorandom number generator (PRNG). As a result, some structural asymmetry (never decreasing even pixels and increasing odd pixels when hiding the data) is introduced, and thus it is very easy to detect the existence of hidden message even at a low embedding rate using some reported steganalytic algorithms, such as the Chi-squared attack [1], regular/singular groups (RS) analysis[2], sample pair analysis [3], and the general framework for structural steganalysis [4]. Generally, the regions located at the sharper edges present more complicated statistical features and are highly dependent on the image contents. To reduce this problem, we propose an edge adaptive scheme and apply it to the LSBMR - based method, high-dimensional image models [5], amplitude of histogram local extrema [6], to improve efficiency LSB matching in image with high - frequency noise[7].Reversible Steganography scheme has the ability to embed the secret data into a host image and then recover the host image without losing any information when the secret data is extracted. This should be overcome by using some techniques. Reversible Steganography is also known as reversible data hiding. No modification is done in the digital representation of the cover image when reversible data hiding method is used. The Reversible data hiding is used in the field of medical, military, legal applications etc.

II. SURVEY

In this section we will be presenting the survey on various data embedding techniques in Steganography to facilitate secure data transmission over the underlying communication network.

A. Edge Adaptive Image Steganography Based on LSB Matching Revisited

Edge Adaptive Image Steganography Based on LSB Matching Revisited scheme introduced by Weiqi Luo.LSBMR[8] uses a pair of pixels as an embedding unit, in which the LSB of the first pixel carries one bit of secret message, and the relationship (odd–even combination) of the two pixel values carries another bit of secret message. LSBMR, deal with each given pixel/pixel pair without considering the difference between the pixel and its neighbors.

(i) Data Embedding

Initializes some parameters, which are used for subsequent data preprocessing and region selection, and then estimates the capacity of those selected regions. If the regions are large enough for hiding the given secret message, then the data hiding is performed on the selected regions. Finally, to obtain the stego image, it does post processing. Otherwise the scheme needs to revise the parameters, and then repeats region selection and capacity estimation until can be embedded completely. May be the parameters are different for different image content and secret message.

(ii) Extracting Data

First extract the side information, i.e., the block size and the threshold from the stego image. As we done exactly the same things in data embedding. The stego image is divided into blocks and the blocks are then rotated by random degrees based on the secret key. The resulting image is rearranged as a row vector . Finally, we get the embedding units by dividing into no overlapping blocks with two consecutive pixels.

Limitations

- Each pixel along the travelling order is dealt separately.
- The secret bit simply overwrites the LSB of the pixel.

B. High – Dimensional Image Models To Perform Highly Undetectable Steganography

Another modified form of data embedding technique is introduced by High – dimensional image scheme Zhang *et al.*Most popular embedding methods used with the digital images is the Least Significant Bit(LSB) replacement, where the LSBs of individual cover elements are replaced with message bits. The modification of LSB Replacement method is LSB Matching(of – ten called ±1 embedding)[11]. This algorithm randomly modulates pixel values by ±1 so that the LSBs of pixels match the communicated message. The use of high – dimensional image models in Steganography that cannot be used in steganalysis Separate the image model from coding, which allows simulating optimal coding and thus comparing the image models without the effect of coding. Moreover the message can be hidden in parts of the image difficult in steganalysis. In high – dimensional image method, when good pixel are selected[14] (e.g.-pixels in noisy and textured areas). When the message is inserted into the image while modifying only the selected pixels by using wet paper codes. This scheme always uses al pixel for embedding, but it changes them with probability inversely proportional to the detectability of their change. Minimizing embedding impact correlated with detectability.

Limitations

- Efficiency of the image is limited while matching pixels.
- The embedding pixels changes with probability of detection.
- Discrimination accuracy is significantly lower than for LSB flipping.

C. The Amplitude Of Histogram Local Extrema

The performance of the amplitude of local extrema compared with the other state – of-the art techniques. The amplitude of local extrema which reduces noise associated with border effects, i.e. pixel values with intensities of either 0 or 255. It describes extension of the amplitudes of local extrema to 2D adjacency histograms. In this techniques which modify the design the ALE steganalyser by incorporating additional complementary features, also based on the amplitude of some local extrema.

(i) Removing Interferences at the Histogram Borders

Embedding based on local extrema introduces a small asymmetry: 0-valued pixels will *always* be changed to 1 if their LSB needs to be modified. Similarly, 255-valued pixels will *always* be changed to 254. This asymmetry in the histogram can cause interferences with the extracted feature, as demonstrated. To avoid this, remove the border effect: where the set of local extrema $E1$ is now reduced to be within [3*,* 252]. In other words, the positions *{*1*,* 2*,* 253*,* 254*}* are simply not considered as potential local extrema.

(ii) Considering 2D Adjacency Histograms

Inspired by [10], the analysis of local extrema has been extended to 2D adjacency histograms, $h2(k, l)$, which tabulate how often each pixel intensity is observed next to another in the horizontal direction:

$$\mathbf{h2}(\boldsymbol{k}, \boldsymbol{l}) =\{(\boldsymbol{i}, \boldsymbol{j}) \in \boldsymbol{I} \mid \mathbf{p}(\boldsymbol{i}, \boldsymbol{j}) = \boldsymbol{k}, \mathbf{p}(\boldsymbol{i}, \boldsymbol{j} + 1)=\boldsymbol{l}\} \quad \mathbf{(1)}$$

where **p**(i, j) is the pixel value at location (i, j) in the input image, and I is a bidimensional index which runs through all pixel locations in the image. Since adjacent pixels have, close intensity values, this histogram is sparse off the diagonal. It should be noted that the histogram defined by Equation (1) can be slightly modified. The asymmetry in the histogram can cause interferences with the extracted feature, as demonstrated. Computing the adjacency histogram instead of the usual intensity histogram.

TABLE 1. TABLE OF ALE FEATURES

1
2
	A1(h1)
	d1(h1)

3	A2(h2) (horizontal direction)
4	A2(h2) (vertical direction)
5	A2(h2) (main diagonal direction)
6	A2(h2) (minor diagonal direction)
7	d2(h2) (horizontal direction)
8	d2(h2) (vertical direction)
9	d2(h2) (main diagonal direction)
10	d2(h2) (minor diagonal direction)

To obtain 3 other adjacency histograms for other directions (vertical, main diagonal, and minor diagonal).

Limitations

- The length of the hidden message is less than the number of pixels in the image.
- High frequency power.

(iii) LSB Matching in image with high – frequency Noise

Jun Zhang introduced LSB Matching in image with high – frequency Noise**.** The histogram of a stego image has less high-frequency power than the corresponding histogram of the cover image. The centre of gravity of the histogram will decrease after LSB matching embedding. While good results were reported on a small test set using color histograms. subsequent experiments revealed that this technique performs poorly on LSB matching in grayscale images. To address this issue, we proposed two novel ways of applying the histogram characteristic function (HCF)1, based on (i) calibrating the output using a down sampled image, and (ii) computing the adjacency histogram instead of the usual intensity histogram. Significant improvements in detection of LSB matching in grayscale images were thereby achieved. detection performance for LSB matching in grayscale images with high levels of high-frequency noise, such as high-resolution scans of photographs. It appears to be very difficult for steganalyzers based on an additive noise model to accurately distinguish between the stego signal and naturally occurring noise in images.

(ii) *Wet Paper Codes with Improved Embedding Efficiency.*

Jessica Fridrich and Miroslav Goljan introduce coding method that empowers the steganographer with the ability to use arbitrary selection channels while substantially decreasing the number of embedding changes, assuming the embedded message length is shorter than 70% of maximal embedding capacity. The method can be flexibly incorporated as a module into majority of existing steganographic methods. Based on syndrome coding using random linear codes , we derive bounds on achievable embedding efficiency for linear codes.

i) Embedding efficiency

By increasing code length, random linear codes asymptotically achieve the theoretical upper bound (6) on the embedding efficiency. However, the computational complexity of the proposed coding method imposes a limit on the practically usable code length. In this section, we derive an approximate but sufficiently accurate expression for the embedding efficiency of the proposed method.

Given two integers p and n, let **H(p, n)** be an ensemble of all binary matrices of dimension **p × n** with n different non-zero columns. The average number of embedding changes for a given matrix **H ∈ H(p, n)** is the average distance Ra to the code represented by H. The efficiency increases with shorter messages for pixels.

Based on syndrome coding using random linear codes , we derive bounds on achievable embedding efficiency for linear codes. Additive noise model to accurately distinguish between the stego signal and naturally occurring noise in images.

TABLE 2 COMPARISON OF DATA EMBEDDING TECHNIQUES

Techniques	**Security**	**Efficiency**	**Access Control**
LSBMR	Mode rate	High	High
High dimensional image	Low	High	High
ALE	High	Low	Low
High frequency noise	Low	Low	Low
Wet paper codes	High	Low	High

III. POSSIBLE SOLUTION

In this survey, we have discussed about various Data Embedding Techniques which conceals data into a carrier for conveying secret messages having the drawback of image distortion. To overcome this problem we proposed a simple and efficient data embedding technique Pixel Pair Matching (PPM) algorithm. The basic idea of PPM is to use the values of Pixel Pair as a reference coordinate and the values of Pixel Pair in the neighborhood set of this Pixel Pair according to a given message digit. The Pixel Pair is then digit. Exploiting Modification Direction (EMD) and Diamond Encoding (DE) are two data-hiding methods proposed recently based on PPM. PPM allows users to select digits in any notational system for data embedding and thus achieves a better image quality.

IV. CONCLUSION

In this paper, we have carried out a wide survey on different approaches for data embedding techniques and analyze various algorithms for data hiding and point out the drawbacks of hiding information. To overcome these techniques, we proposed a simple and efficient data embedding method based on Pixel Pair Matching (PPM) algorithm. PPM not only resolves the low – payload problem but also offers smaller Mean Square Error (MSE) compared with Optimal Pixel Adjustment Process and Diamond Encoding.

REFERENCES

[1] A.Westfeld and A. Pfitzmann, "Attacks on steganographic systems, "in *Proc. 3rd Int. Workshop on Information Hiding*, 1999, vol. 1768, pp. 61–76.

[2] J. Fridrich, M. Goljan, and R. Du, "Detecting LSB Steganography in color, and gray-scale images," *IEEE Multimedia*, vol. 8, no. 4, pp. 22–28, Oct. 2001.

[3] S. Dumitrescu, X.Wu, and Z.Wang, "Detection of LSB Steganography via sample pair analysis," *IEEE Trans. Signal Process.*, vol. 51, no. 7, pp. 1995–2007, Jul. 2003.

[4] A. D. Ker, “A general framework for structural steganalysis of LSB replacement,” in *Proc. 7th Int. Workshop on Information Hiding*, 2005, vol. 3427, pp. 296–311.

[5] A. D. Ker, “A fusion of maximum likelihood and structural steganalysis,” in *Proc. 9th Int. Workshop on Information Hiding*, 2007, vol.4567, pp. 204–219.

[6] J. Fridrich*, Steganography in Digital Media: Principles, Algorithms, and Applications*, 1st ed. Cambridge, U.K.: Cambridge Univ. Press,2009.

[7] G. Cancelli, G. Doerr, M. Barni, and I. Cox, “A comparative study of steganalyzers,” in *Proc. IEEE Workshop Multimedia Signal Process.*, 2008, pp. 791–796.

[8] J. Zhang, I. Cox, and G. Doerr, “Steganalysis for LSB matching in images with high-frequency noise,” in *Proc. IEEE Workshop Multimedia Signal Process.*, 2007, pp. 385–388.

[9] J. Fridrich, M. Goljan, and D. Soukal, “Wet paper codes with improved embedding efficiency,” *IEEE Trans. Inf. Forensics Security*, vol. 1, no.1, pp. 102–110, Mar. 2006.

[10] C.-H. Yang, C.-Y. Weng, S.-J. Wang, and H.-M. Sun, “Adaptive data hiding in edge areas of images with spatial LSB domain systems,”*IEEE Trans. Inf. Forensics Security*, vol. 3, no. 3, pp. 488–497, Sep. 2008.

[11] T. Pevný, T. Filler, and P. Bas, “Using high-dimensional image models to perform highly undetectable Steganography,” in *Information Hiding*. New York, NY, USA: Springer, 2010, vol. 6387, LNCS, pp. 161–177.

[12] G. Cancelli, G. Doerr, I. Cox, and M. Barni, “Detection of LSB Steganography based on the amplitude of histogram local extrema,” in *Proc. IEEE Int. Conf. Image Process.*, 2008, pp. 1288–1291.

[13] S. Lyu and H. Farid, “Steganalysis using higher-order image statistics,” *IEEE Trans. Inf. Forensics Security*, vol. 1, no. 1, pp. 111–119, Mar.2006.

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