http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 9 (8): 1725-1738, 2010
ISSN 1812-5638

A Review on Detection of LSB Matching Steganography

[1]Jiaohua Qin, [2]Xuyu Xiang and [2]Meng Xian Wang
[1]School of Computer and Information Engineering, Central South University of Forestry and Technology, Changsha, 410004, China
[2]Department of Mathematics and Computer, Hunan City University, Yiyang, 413000, China

Abstract: LSB matching steganalysis techniques detect the existence of secret messages embedded by LSB matching steganorgaphy in digital media. This study presents a survey of LSB matching steganalysis methods for digital images. Firstly, study described the structure of LSB matching steganalysis, which includes three parts: LSB matching steganography, detectors for LSB matching and the evaluation methodology. Secondly, study classified the existing detection algorithms into two categories according to the fact that the main contribution of the algorithm is detector or estimator. For the detectors, study classified the existing various methods to two categories, described briefly their principles and introduced their detailed algorithms. For the estimators, study introduced the existing two estimating methods for LSB matching. Finally, study concluded and discussed some important problems in this field and indicated some interesting directions that may be worth researching in the future.

Key words: Detector, estimator, LSB matching, steganalysis, steganography

INTRODUCTION

The goal of steganography is to hide the very presence of communication by embedding messages into innocuous-looking cover objects (Fridrich *et al*., 2005). The most popular, frequently used and easy to implement steganographic method is the Least Significant Bit (LSB) steganography. The LSB steganographic methods can be classified into the following two categories: LSB replacement and LSB matching (also named ±1 embedding) (Melikainen, 2006).

Both LSB replacement and ±1 embedding select a subset of the pixels pseudorandomly using a secret key known to both sender and receiver. In LSB replacement, the least significant bit of each selected pixel is replaced by a bit from the hidden message. And the even pixel values are either unmodified or increased by one, while odd ones are either decreased by one or left unchanged. Note, on average only half these bits will actually be changed; for the other half, the message bit is the same as the image bit already there. This imbalance in the embedding distortion was recently utilized to detect secret messages. There is now substantial literature on LSB replacement such as (Fridrich *et al*., 2001; Dumitrescu *et al*., 2003; Ker, 2004a, b; Jiao-Hua *et al*., 2007a, b; Niu *et al*., 2009) describing sensitive statistical methods for its reliable detection.

As a counter-technology of steganography, steganalysis is a kind of art and science of revealing the secret messages. The steganalysis can disclose drawbacks of steganographic schemes by proving that a secret message has been embedded in a cover, on the other hand, it can prevent the utilization of outstanding steganographic methods by criminals to unlawfully transmit nocuous messages.

The LSB matching, a counterpart of LSB replacement, retains the favourable characteristics of LSB replacement, it is more difficult to detect from statistical perspective. In LSB matching, if the bit must change, the operation of ±1 is applied to the pixel value. The use of + or - is chosen randomly and has no effect on the hidden message. The detectors for both LSB replacement and ±1 embedding work the same way: the LSB for each selected pixel is the hidden bit. Since LSB techniques are fairly easy to implement and have a potentially large payload capacity, there is a large selection of steganography software available for purchase and via shareware (e.g., www.stegoarchive.com). This seemingly innocent modification of the LSB embedding is significantly harder to detect, because the pixel values are no longer paired. Theoretical analysis and practical experiments show that steganalysis of LSB matching is more difficult than that of LSB replacing (Ker, 2005a). As a result, none of the existing attack methods on LSB replacement can be adapted to attack LSB matching.

Harmsen and Pearlman (2003) proposed a steganalysis method using the Histogram Characteristic Function (HCF) as a feature to distinguish the cover and

Corresponding Author: Jiaohua Qin, School of Computer and Information Engineering, Central South University of Forestry and Technology, Shaoshan South Road, No. 498, Changsha, Hunan, 410004, P.R. China Tel: 86-13517490245

stego images. This method is efficient in detecting the LSB replacement for RGB color bitmaps, but ineffective in detecting the LSB matching for grayscale images. Ker (2005a) extended Harmsen's method by two novel ways: (1) calibrating the output center of mass (COM) using a down sampled image, (2) computing the adjacency histogram instead of the usual histogram. Significant improvements in detection of LSB matching in grayscale images were thereby achieved. Yu and Babaguchi (2008b) also extended HCF and used the fusion of the COM of the run-length HCF and Ker's two-dimensional adjacency histogram to detect the LSB Matching. Zhang *et al.* (2007) proposed a method for steganalysis of LSB Matching in images with high-frequency noise. This method has superior results when the images contain high-frequency noise, e.g. uncompressed imagery such as high-resolution scans of photographs and video. However, the method is inferior to the prior art only when applied to decompressed images with little or no high-frequency noise. Fridrich *et al.* (2005) proposed a maximum likelihood estimator for estimating the number of embedding changes for non-adaptive ±K embedding in images. However, they observe that this approach is not effective for never-compressed images derived from a scanner.

There also exist blind techniques such as (Holotyak *et al.*, 2005b; Goljan *et al.*, 2006; Lyu and Farid, 2004), which are some what effective, but they have poor detection performance for LSB matching in grayscale images. Farid (2002) first proposed a framework for learning-based steganalysis and demonstrated it as an effective approach to cope with the steganalysis difficulties caused by various image textures and unknown steganography algorithms. Subsequently, some works have been developed which based on all kinds of features extracted from different domains such as spatial domain (Avcibas *et al.*, 2003), DCT domain (Chen and Shi, 2008; Shi *et al.*, 2006; Pevni and Fridrich, 2007; Xia *et al.*, 2010) and DWT domain (Farid and Lyu, 2003; Xuan *et al.*, 2005; Jiao-Hua *et al.*, 2007a) etc. Luo *et al.*(2008) gave a survey on blind detection for image steganography. However, researches show that the improved performance of image steganalysis is achieved at the expense of increasing the number of the features. Some works mentioned the reduction of feature number utilizing SFFS (Wang and Moulin, 2007), SFS (Miche *et al.*, 2006), PCA (Holotyak *et al.*, 2005a), Hybrid Genetic Algorithm (Xia *et al.*, 2009a, b) and PFSP (Qin *et al.*, 2009a, b).

In this study, we gave an overview of the detection methods for LSB matching steganography. To begin with, we described the structure of LSB matching steganalysis, which includes three parts, namely, LSB matching steganography, detectors for LSB matching and the evaluation methodology. Then we classified the existing detection algorithms into two categories according to the fact that the main contribution of the algorithm is detector or estimator. For the detectors, we classified the existing various methods to two categories, described briefly their principles and introduced their detailed algorithms. For the estimators, we introduce the existing two estimating methods of LSB matching. At last, some important problems in this field are concluded and discussed and some interesting directions that may be worth researching in the future are indicated.

STRUCTURE OF LSB MATCHING STEGANALYSIS

LSB matching steganography: Least Significant Bit (LSB) matching steganography, also named ±1 embedding, is a slightly more sophisticated version of LSB embedding. A grayscale n×m image will be represented with a two-dimensional array of integers x_{ij}, $x_{ij} \in \{1,...,255\}$, $I \in \{1,..., m\}$, $j \in \{l,..,m\}$. A true color 24 n×m bit image will be represented as three grayscale n×m images r_{ij}, g_{ij}, b_{ij}. The distortion due to non-adaptive LSB matching is modeled as an additive i.i.d. noise signal η with the following Probability Density Function (PDF) with $\rho \in [0,1]$ (Fridrich *et al.*, 2005):

$$\begin{cases} P(\eta = 0) = 1 - \frac{\rho}{2} \\ P(\eta = 1) = P(\eta = -1) = \frac{\rho}{4} \end{cases} \quad (1)$$

where, ρ is the embedding rate, that is the ratio between the size of the LSB plane and the length of the message. The LSB matching operation can be described as Table 1.

Detectors for LSB matching: A series of steganalyzers have been developed for LSB Matching Steganography. They can be roughly considered as sharing a common architecture, namely (1) feature extraction in some domain and (2) Fisher Linear Discriminant (FLD) analysis to obtain a 2-class classifier (Cancelli *et al.*, 2008). We consider some possible detectors for LSB Matching, include those of Westfeld's Detector (Westfeld, 2002), Harmsen's HCF COM Detectors (Harmsen and Pearlman, 2003; Harmsen *et al.*, 2004), Ker's extended HCF COM Detectors (Ker, 2005b), Ker's IHCF COM Detectors (Ker, 2005b) and Abolghasemi's Co-Occurrence Matrix (Abolghasemi *et al.*, 2008) and so on.

Table 1: LSB matching operation

	To embed bit b, modify x to	
Pixel value x	b = 0	b = 1
2i, 0<2i<255	2i	2i+1 or 2i-1
2i+1, 0<2i+1<127	2i or 2i+2	2i+1
0	0	1
255	254	255

Other Estimators for steganography, applicable to LSB Matching, include those of Fridrich's Maximum Likelihood Estimator (Fridrich *et al*., 2005) and so on. Those detectors and estimators are briefly reviewed in the next sections.

Evaluation methodology: It is important to have confidence in steganography detectors. A detector is a discriminating statistic, a function of images which takes certain values in the case of stego images and other values in the case of innocent cover images.

Assuming that a detector aims only to give a binary diagnosis of steganography or no steganography and that the detection statistic is a one-dimensional†, the reliability is given by the Receiver Operating Characteristic (ROC) curve, which shows how the false positives and false negatives vary as the detection threshold is adjusted. Because there are a number of steganalysis algorithms we wish to test, each with a number of possible variations, a number of hidden message lengths and tens of thousands of cover images, there are millions of calculations to perform. To do so quickly, we use a small distributed network to undertake the computations; each node runs a highly-optimised program dedicated to the simulation of steganographic embedding and the computation of many different types of detection statistic; the calculations are queued and results recorded, in a database from which ROC curves can be extracted and graphed. This distributed system has been used to analyse the detection of both LSB Replacement and LSB Matching steganography.

In practice, the performance of steganalysis methods is highly dependent on the types of cover images used.

DEVELOPMENT OF DETECTORS FOR LSB MATCHING

One of the earliest detectors suggested for LSB Matching is due to Westfeld, which is based on close colour pairs (Westfeld, 2002).

Westfeld's detector: Westfeld's detector is only applicable to colour images. It is founded on the assumption that cover images contain a relatively small number of different colours, in a very similar way to an early detector for LSB Replacement due to Fridrich *et al*. (2000). Consider a pixel colour as a triple (r, g, b), specifying the red, green and blue components. Fridrich *et al* . (2000) considered colour pairs to detect LSB encoding in colour images. Two colours (r_1, g_1, b_1) and (r_2, g_2, b_2) are a close pair, if $|r_1 - r_2| \leq 1$, $|g_1-g_2| \leq 1$ and $|b_1-b_2| \leq 1$. Westfeld calls these pairs neighbours. The LSB

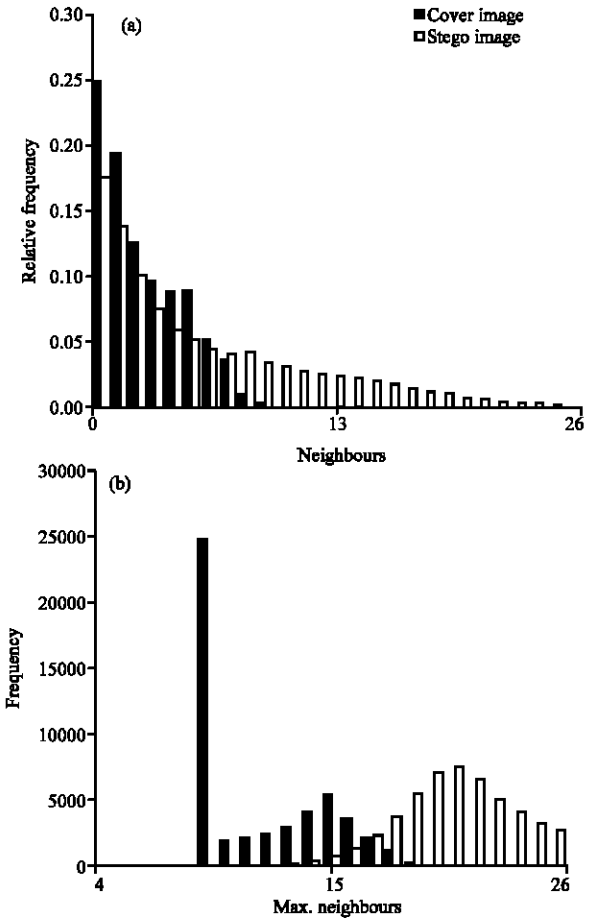

Fig. 1: (a) Relative frequency of the number of neighbours, before and after embedding a maximal-length message by LSB Matching. (b) Histogram of the maximum neighbours statistic, before and after embedding a message 1% of maximal length

Matching algorithm will turn a large number occurrences of a single colour into a cluster of closely-related colours. Each colour can have up to 26 neighbours (excluding itself). A colour in a carrier medium has only 4 or 5 neighbours on average and that, in JPEG images, no colour has more than 9 neighbours. On the other hand, after embedding a message using LSB Matching (even when the message is quite small) enough new colours are created that the average number of neighbours is substantially increased and many colours even have the full complement of 26 neighbours. For example, Fig. 1a. The number of neighbours of each colour in a JPEG image has been computed and the histogram displayed. The average number of neighbours for each colour is 2.20. This is repeated after embedding a maximal-length random message (3 bits per cover pixel) by LSB Matching; the average is now 5.58.

The detector remains perfect for JPEG images by using the histogram of the maximum neighbours statistic. Even when a message only 1% of the maximum is embedded the detector still functions very well. But the story is quite different for cover images which are not JPEGs. In particular, it is false for JPEG images which have been even slightly modified by image processing operations such as re-sizing, because that each colour has a number of its possible neighbours occurring in the cover image.

Histogram characteristic function detectors: Some detectors for LSB Matching in the literature is due to (Harmsen and Pearlman, 2003; Harmsen *et al*., 2004). In fact, Harmsen's detector is designed to work on any type of steganography which can be modelled as additive noise. It is clear that LSB Matching is one such type.

Harmsen's HCF COM detectors: Harmsen calls $H_c[k]$ the Histogram Characteristic Function (HCF) of the N-element DFTs for the histogram $h_c(n)$ of cover image. They consider that the steganographic embedding can be modeled as independent additive noise. Therefore, the addition of integer random variables corresponds to the convolution of their mass functions, $h_{s^-} = h_c * f_\Delta$. The distribution of the added noise in the case of LSB Matching, when the hidden message is of maximal length, is just:

$$f_\Delta(n) = \begin{cases} 0.25 & n = -1 \\ 0.5 & n = 0 \\ 0.25 & n = 1 \end{cases} \quad (2)$$

Let $H_C[k]$, $H_S[k]$ and $F_\Delta[k]$ be the N-element discrete Fourier transforms(DFTs) of $h_c(n)$, $h_s(n)$ and $f_\Delta(n)$, respectively. Elementary calculation gives that $F_?(k) = \cos^2(?k/N)$; this monotone function, always no greater than 1, drops to zero as k reaches N/2. Therefore, $H_S[k]$ will be no larger than $H_C(k)$ and for large k will be appreciably smaller. Then, we have:

$$H_S(k) = H_C(k)F_\Delta(k) \leq H_C(k) \quad (3)$$

To diagnose the presence of steganography Harmsen uses the centre of mass (COM) of the HCF:

$$C(H[k]) = \sum_{i=0}^{N/2} i\ H[i] \Big/ \sum_{i=0}^{N/2} H[i] \quad (4)$$

After steganographic embedding:

$$C(H_S[k]) \leq C(H_C[k]) \quad (5)$$

It is simply this COM that is the discriminator for detecting steganogtaphy in RGB color images. However, there are two potential weaknesses in HCF COM detectors. The first is that the value of the HCF COM is essentially without context – it is difficult to say whether a particular value is "low" or "high" as it may depend as much on the type of cover image than the presence or absence of steganography. The second is that the HCF COM depends only on the histogram of the image and so is throwing away a great deal of structure.

Ker (2005b) pointed out that Harmsen's HCF COM detector performed poorly indeed, especially for gray scale images. Ker extended Harmsen's method on LSB matching.

Ker's extended HCF COM detectors: Ker consider, the values of $C(H_C[k])$ depend heavily on the source of cover image and worse, there is high variability amongst $C(H_C[k])$, often far more than the typical differences between $C(H_C[k])$ and $C(H_S[k])$ (Fig. 2). The significant weakness of this method is that the detector does not see the cover image and so does not know $C(H_C[k])$.

By calibrating the output COM using a down-sampled image and computing the adjacency histogram instead of the usual histogram, Ker proposed his new method on uncompressed grayscale images.

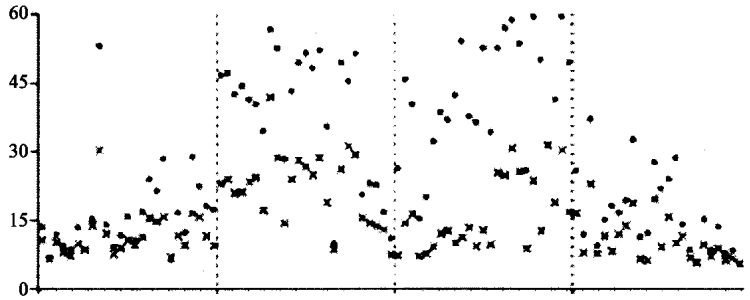

Fig. 2: Values of C(H[k]) (circles) before and (crosses) after embedding from four different sources

Ker's calibrated HCF COM detector: Consider downsampling an image by a factor of two in both dimensions using a straightforward averaging filter. Precisely, let $p_c(i, j)$ be the pixel intensities of the downsampled cover image given by:

$$p_c'(i,j) = \left\lfloor \sum_{u=0}^{1}\sum_{v=0}^{1} \frac{p_c(2i+u, 2j+v)}{4} \right\rfloor \quad (6)$$

and $p_s(i, j)$ the similarly downsampled version of the stego image. They divide the summed pixel intensities by four and take the integer part to reach images with the same range of values as the originals. By computing the HCF and COM of these two downsampled images $C(H_C[k])$ and $C(H_S[k])$, Ker use C(H[k]/C(H[k]) as a dimensionless discriminator.

Ker's adjacency HCF COM detector: The procedure of adjacency histogram method is very similar to the procedure of calibration method. One difference is that the two-dimensional adjacency histogram is defined as fallows:

$$h^{(2)}(m,n) = \left|\{(i,j) \mid p(i,j) = m, p(i,j+1) = n\}\right| \quad (7)$$

As before, we form the HCF by using a two dimensional DFT and the two-dimensional COM.

The detection performance of Ker's Detectors are given by using receiver operating characteristic (ROC) curves and shown in Fig. 3.

The Fig. 3a is generated from 20000 images that have been subject to fairly harsh JPEG compression and the Fig. 3b is from 3000 uncompressed bitmaps.

From Fig. 3 can see that the detection in JPEG compressed covers becomes extremely reliable by using Ker's extended detectors and significant improvements in detection of LSB matching in grayscale images were thereby achieved. However, the detector degrades gracefully with shorter messages.

Ker's IHCF COM detectors: As above, Ker gave two methods to improve HCF COM detectors for grayscale images. Then, Ker (2005b) expand his recently-developed techniques for the detection of LSB Matching in grayscale images into the full-colour case.

The obvious alternative is not to do any dividing or rounding; in this case we are not downsampling and so we might as well consider pixels in pairs rather than groups of 4.

Ker consider a squeezed version of the original image as follows:

$$p_{C'}(i,j) = p_C(2i, j) + p_C(2i+1, j) \quad (8)$$

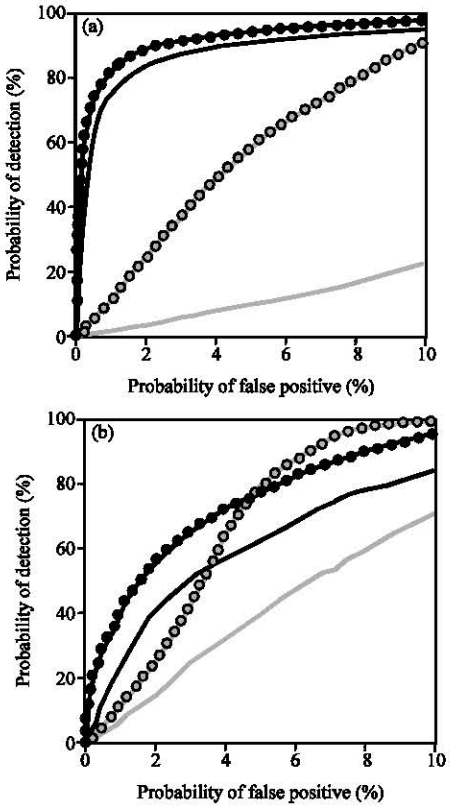

Fig. 3: ROC curves

This detector is, in most cases, a large step up in sensitivity from the others discussed here. Experimental results show the extent of this, in the case of one set of JPEG images, even after the JPEGs are resampled. It can detect steganography with reasonable reliability even when the hidden message is only 5% of the maximum.

Yu's RLHCF COM detectors

RLHCF COM detector: Yu and Babaguchi (2008a) calculate and analyze the run length histogram. They find that run length histogram can be used to define a feature such as HCF. They call this feature run length histogram characteristic function (RLHCF) and use the center of mass (COM) of the RLHCF

$$C(\vec{h}) = \frac{\sum_{j=1}^{n} jh(j)}{\sum_{j=1}^{n} h(j)} \quad (9)$$

where, n is the maximum run length. Because of the shrinking effect of run length histogram after embedding, there is $C(\vec{h}_s) < C(\vec{h}_c)$. They calculate the alteration rate R by using

$$R = \frac{C(\vec{h}) - C(\vec{h}^*)}{C(\vec{h})}$$

and calculate the HCF COM $C^2(H^2[k, 1])$ using Ker's method, then normalize R and $C^2H^2[k, 1]$ to a common range [0,1].

Comparing the value $C^2(H^2[k, 1])+R$ with a predetermined threshold, it can distinguish the stego images from cover images.

From Fig. 4, we can see, the RLHCF COM detector is reliable than Ker's detector.

Fusion extended HCF and RLHCF COM detector: Yu and Babaguchi (2008a) further extend the COM to high order as features for steganalysis. The nth statistical moment (n-th COM) of HCF is defined:

$$C_n^{DFT}(H[k]) = \sum_{i=1}^{N/2} i^n |H[i]| \Big/ \sum_{i=1}^{N/2} |H[i]| \quad (10)$$

$$C_n^{DFT}(H[k,l]) = \sum_{i,j=0}^{N/2} (i+j)^n |H[i,j]| \Big/ \sum_{i,j=0}^{N/2} |H[i,j]| \quad (11)$$

As before, they form the extended HCF and RLHCF COM detector by comparing the value

$$F = \sum_{n=1}^{3} (C_n^{DFT}(H[k]) + C_n^{DFT}(H[k,l]) + R_n)$$

with a predetermined threshold, so can determine whether the given image is a stego image.

Xia's NDHCF COM detector: Hu *et al.* (2008) adopt image segmentation to separate image into different domains and analyze the statistic property of node degree for Minimum Spanning Tree (MST) in random domain. Xia *et al.* (2009a, b) propose a method to detect Least Significant Bit (LSB) matching steganography which is based on neighbourhood Node Degree Histogram Characteristic Function (NDHCF). First we calculate the center of mass (COM) of the NDHCF then embed another random secret message to compute the alteration rate *R* of the NDHCF COM.

The neighbourhood node degree of p(i,j) is defined as following:

$$d(i,j) = |\{(i+u, j+v) \mid p(i+u, j+v) = p(i,j)\}| \quad (12)$$

The neighborhood node degree histogram(NDH) is defined as following:

$$h(x) = |\{(i,j) \mid d(i,j) = x\}| \quad (13)$$

we use the center of mass of the neighborhood node degree histogram C(h(x)) and two-dimensional NDHCF COM $C^2(h^2(x, y))$.

We select NDHCF COM and the alteration rate as features and use support vector machines as a classifier. For a given image, we compute the features (C(h(x)), R, $C^2(h^2(x, y))$ and R^2) twice using 3×3 and 5×5 neighborhood respectively, which form an 8-D feature vector for steganalysis. Experimental results demonstrate (Fig. 5) that the proposed method is efficient to detect the LSB matching stegonagraphy on compressed or uncompressed images.

From the experiment results, we can see that the NDHCF COM Detector outperforms other three methods. Under the same probability of false positive, the detection rate of our method is much higher than others.

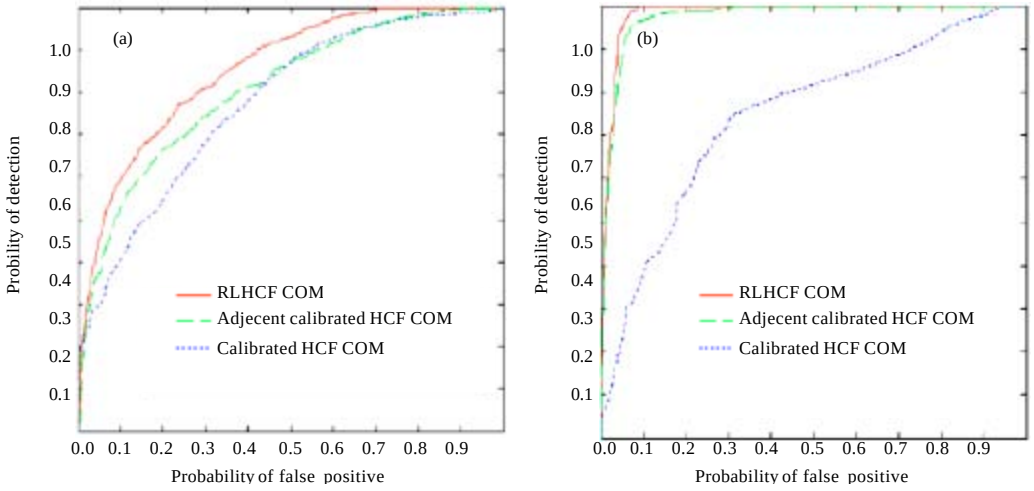

Fig. 4: ROC curves Compared with Ker's method for (a) uncompressed images and (b) JPEG images

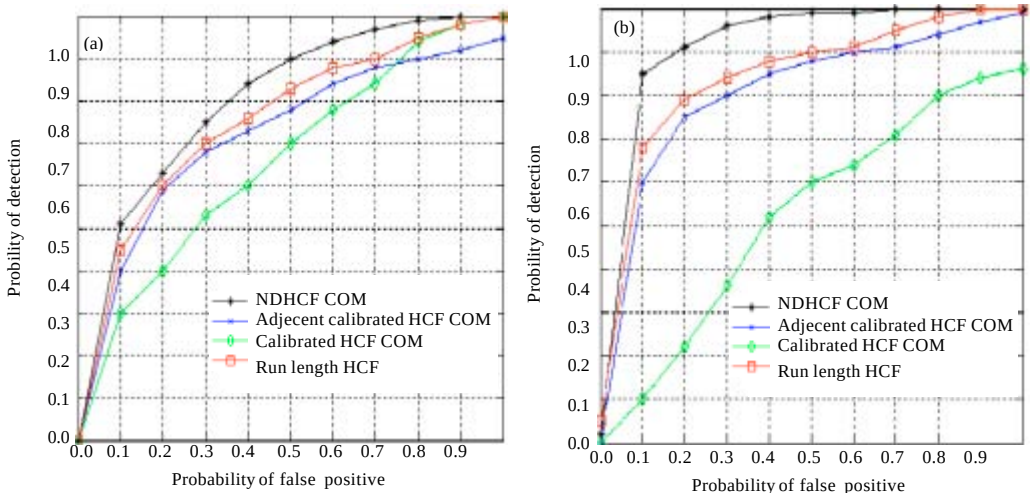

Fig. 5: ROC curves for (a) uncompressed images and (b) compressed images

Feature mining detectors

Zhang's ALE(Amplitude of local extrema) Detector: Zhang *et al.* (2007) consider the sum of absolute differences between each local extremum and its neighbors in the histogram. These sums are denoted Dc and Ds for the cover and stego images, respectively. That is:

$$D_c = \sum_{n^*} |2.h_c(n^*) - h_c(n^*-1) - h_c(n^*+1)| \quad (14)$$

$$D_s = \sum_{n^*} |2.h_s(n^*) - h_s(n^*-1) - h_s(n^*+1)| \quad (15)$$

where, n is local minimum.

For any image after LSB matching steganography, it has Dc>Ds.

Figure 8 demonstrates a significant improvement in performance over that of Ker (2005b) and GFH (Goljan *et al*., 2006).

The experimental results demonstrate that the histogram extrema method has substantially better performance. However, if the datasets are JPEG compressed with a quality factor of 80, the high frequency noise is removed and the histogram extrema method performs worse.

Qin's DNPs and DLENs detector: A novel steganalysis method, which exploits the difference statistics of neighboring pixels, is proposed by Qin *et al.* (2009a) to detect the presence of spatial LSB matching steganography. In this method, the differences between the neighboring pixels (DNPs), the differences between the local extrema (DLENs) and their neighbors in grayscale histogram are used as distinguishing features and the SVM is adopted to construct classifier.

The difference histogram of images on horizontal is defined as follows:

$$H_1(d) = \frac{\sum_{i=0}^{M-1}\sum_{j=0}^{N-2}\delta(|p_{i,j} - p_{i,j+1}|, d)}{M \times (N-1)} \quad (16)$$

The sums of DNPs with the value of zero and that with the value larger than one are denoted as F_1 and F_2, respectively.

$$F_1 = \sum_{i=1}^{4} H_i(0) \quad (17)$$

$$F_2 = \sum_{i=1}^{4}\sum_{d=2}^{d_{max}} H_i(d) \quad (18)$$

where, i = 1,2,3,4 means the direction of horizontal, vertical, 45 and 135 degree diagonal.

The sum of the absolute differences between the local maximums and their neighbours in a cover image histogram is denoted as S_{max}.

$$S_{max} = \sum_{x^*_{max} \in h_c} |2h_c(x^*_{max}) - h_c(x^*_{max}-1) - h_c(x^*_{max}+1)| \quad (19)$$

The sum of the absolute differences between $h_s(x^*_{max})$ and their neighbours is given by:

$$S^*_{max} = \sum_{x^*_{max} \in h_c} |2h_s(x^*_{max}) - h_s(x^*_{max}-1) - h_s(x^*_{max}+1)| \quad (20)$$

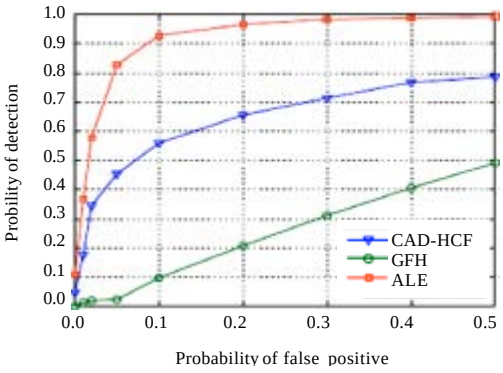

Fig. 6: ROC curves with an embedding rate of ρ = 0.5

Similarly, we denote the sum of absolute differences between the local minimums and their neighbours in a cover image histogram as S_{min} and denote the absolute differences between $h_s(x^*_{min})$ and their neighbours as s^*_{min}.

The change rate of the feature F_i before and after LSB matching steganography is denoted as:

$$R_i = \frac{F_i - F_i^*}{F_i} \quad (21)$$

For a given image, we compute the features (F_1, F_2, S_{max}, S_{min} and their change rate) to form an 8-D feature vector for steganalysis. Experimental results show (Fig. 6) that the proposed method is efficient to detect the LSB matching steganography for the compressed and uncompressed images and outperforms other recently proposed algorithms.

From Fig. 7, it is easy to see that this method achieves higher detection accuracy than the previous methods do. And both for the compressed images and the uncompressed images, this method can obtain better performance.

Huang's neighbourhood gray levels detector: For a given image, Huang *et al.* (2007) get an image by combining the least two significant bit-planes and divide it into 3×3 overlapped subimages. According to the count of comprised gray levels, these obtained subimages are grouped into four types, i.e., T_1, T_2, T_3 and T_4, where T1 includes the subimages in which all the pixels have the same value. Through embedding a random sequence by LSB matching and computing the alteration rate of the number of elements in T1, they find that normally the alteration rate is higher in cover image than the value in the corresponding stego image, which is used as the discrimination rule in their detector.

Suppose an M×N grayscale image I(x, y) is composed of eight 1-bit planes I_0~I_1, ranging from

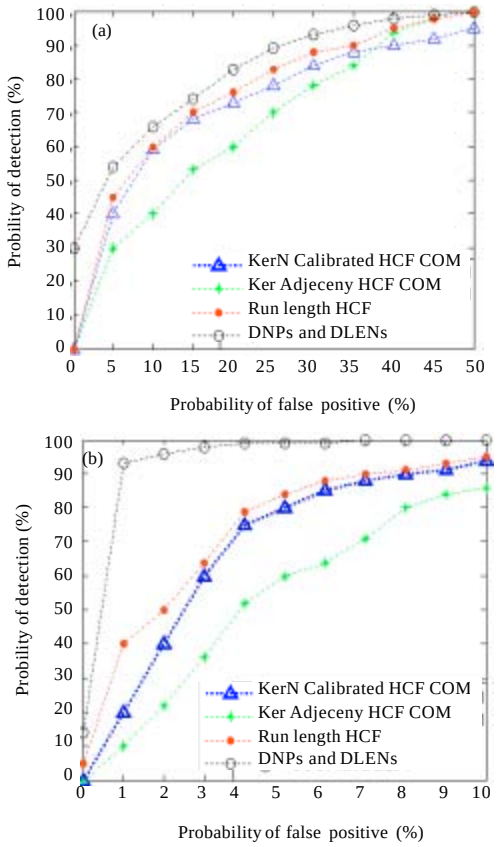

Fig. 7: ROC curves for (a) uncompressed images and (b) compressed images

bit-plane 0 for the least significant bit to bit-plane 7 for most significant bit. We get an image A(*x*,*y*) by combining the least two significant bit-planes as follows:

$$A(x,y) = I_0(x,y) + I_1(x,y) \times 2 \quad (22)$$

The alteration rate *k* are obtained by using:

$$k = \frac{|T_1| - |T_1^*|}{|T_1|} \quad (23)$$

where, $|T_1|$ denotes the number of elements belonging to T_1.

Comparing the value k with a predetermined threshold, it can determine whether the given image is a stego image.

They compare the method with Ker's two methods, the stego images with the secret message length p = 1. The experimental results show that, under the same probability of false positive, the detection rate of our method is much higher than Ker's two methods. However,

if the stego image contains too small amount of hidden data compared with the carrier image size and thus no secret message bit has been embedded into the 5×5 sub region, it is difficult for us to distinguish the cover and stego images using this detector as a discrimination rule.

Liu's feature mining detectors:

Liu's CF detector: Liu *et al.* (2005, 2006) indicate that the significance of features and the detection performance depend not only on the information-hiding ratio, but also on the image complexity. They(Image complexity and feature mining for steganalysis of least significant bit matching steganography (Liu *et al.*, 2008) introduce a parameter of image complexity that is measured by the shape parameters of the Generalized Gaussian Distribution (GGD) in the wavelet domain and use the Correlation Features(CF) to design the detector.

They consider the correlation between LSBP and the second least significant bit plane (LSBP2). M1(1:m,1:n) denotes the binary bits of the LSBP and M2(1:m,1:n) denotes the binary bits of the LSBP2.

The covariance function is defined as:

$$\mathrm{Cov}(x_1,x_2)=\mathrm{E}[(x_1-u_1)(x_2-u_2)] \tag{24}$$

where, ui = $\mathrm{E}(x_i)$.

C1 is defined as follows:

$$C_1=\mathrm{cor}(M_1,M_2)=\frac{\mathrm{Cov}(M_1,M_2)}{\sigma M_1\sigma M_2} \tag{25}$$

The autocorrelation C(k, l) of the LSBP is defined as follows:

$$C(k,l)=\mathrm{cor}(X_k,X_l) \tag{26}$$

where, $X_k=M_1(1:m-k,1:n-l)$; $X_l=M_1(k+1:m,l+1:n)$. Setting k and l to different values, the features from C_2 to C_{15} are presented.

The autocorrelation coefficients C_{16} and $C_H(l)$ are defined as:

$$C_{16}=\mathrm{cor}(H_e,H_o) \tag{27}$$

$$C_H(l)=\mathrm{cor}(H_{l1},H_{l2}) \tag{28}$$

where, He, Ho, Hl1 and Hl2 are the histogram probability densities.

Set l = 1, 2, 3 and 4; the features C_{17}-C_{20} are obtained. The correlation features in the difference domain are given as follows:

$$C_E(t;k,l)=\mathrm{cor}(E_{t,k},E_{t,l}) \tag{29}$$

where, $E_{t,k}=E_t(1:m-k,1:n-l)$; $E_{t,l}=E_t(k+1:m,l+1:n)$. Setting different values to *t*, *k* and *l*, features C_{21}-C_{41} are obtained.

The experiments show that the statistical significance of features and the detection performance closely depend, not only on the information-hiding ratio, but also on the image complexity. While, the hiding ratio decreases and the image complexity increases, the significance and detection performance decrease. Meanwhile, the steganalysis of LSB matching steganography in grayscale images is still very challenging in the case of complicated textures or low hiding ratios.

Liu's EHPCC detector: To improve the performance in detecting LSB matching steganography in grayscale images, based on the previous work (Image complexity and feature mining for steganalysis of least significant bit matching steganography (Liu *et al.*, 2008) propose five types of features EHPCC (Entropy, high-order statistics, probabilities of the equal neighbors, correlation features, complexity) and introduce a dynamic evolving neural fuzzy inference system (DENFIS).

The entropy of NNH (NNH_E) is calculated as follows:

$$\mathrm{NNH\_E}=-\sum\rho_H\log_2\rho_H \tag{30}$$

The rth high-order statistics of NNH (NNH_HOS) is given as:

$$\mathrm{NNH\_HOS}(r)=\frac{(1/N^3)\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}\sum_{z=0}^{N-1}(H(x,y,z)-(1/N^3)\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}\sum_{z=0}^{N-1}H(x,y,z))^r}{\sigma_H^r} \tag{31}$$

where, NNH denotes the distribution density of the NNH.

- Shape parameter β of the GGD of the HH wavelet sub-band that measures the image complexity
- Entropy of the histogram of the nearest neighbors, NNH_E
- The high-order statistics of the histogram of the nearest neighbors, NNH_HOS(r) and r is set from 3 to 22, total 20 high-order statistics
- The Probabilities of the equal neighbors(PEN), include the correlation between the Least Significant Bit Plane (LSBP) and the second least significant bit Plane (LSBP2) and the correlation in the LSBP and the autocorrelation in the image histogram; The correlation in the difference between the image and the denoised version

- Correlations features consist of C1, C(k, l), C2, CH(l) and CE(t ; k, l), described in Liu's CF Detector

By seting the following lag distance to k and 1 in C(k,l) and 14 features are obtained:

- k = 0, l = 1, 2, 3 and 4; l = 0, k = 1, 2, 3 and 4.
- k = 1, l = 1; k = 2, l = 2; k = 3, l = 3; k = 4 and l = 4
- k = 1, l = 2; k = 2, l = 1

The experimental results also indicate that image complexity is an important parameter to evaluation of the detection performance. At a certain information-hiding ratio, it is much more difficult to detect the information-hiding behavior in highimage complexity than that in low complexity.

Abolghasemi's co-occurrence matrix detectors: Considering the asymmetry of the co-occurrence matrix, Abolghasemi *et al.* (2008) adopted the elements of the main diagonal and a part of the upper and lower of main diagonal from co-occurrence matrix, as shown in Fig. 8, to construct the feature vector. We reshape diagonal elements of co-occurrence matrix as following:

$$F = \{D_{-256}, \cdots, D_{-2}, D_{-1}, D_0, D_1, D_2, \cdots, D_{256}\} \quad (32)$$

In the experimental work, for cases 3 Bp, 4 Bp and 5 Bp (Fig. 9) they consider whole of elements of F as feature vector and for the cases more than 5 bit planes only consider feature vector as following:

$$F = \{D_{-2}, D_{-1}, D_0, D_1, D_2\} \quad (33)$$

This method extract features from cooccurrence matrix of an image which some of its most significant bit planes are removed. The experimental results indicate, for the LSB Matching embedding it is shown that by removing 3 significant bit planes detection rates were increased.

Marvel's bitplane-CTW (context tree weighting) detector: Boncelet and Marvel (2007) use a lossless compression technique to compress the last two bitplanes in an effort to model the image structure where the data may be hidden. The lossless compression we use is called BCTW, for Bitplane-CTW, where CTW is the Context TreeWeighting algorithm. BCTW compresses an image bitplane by bitplane, from the most significant to the least significant. BCTW uses two different contexts, one for the most significant bitplane and one for all other bitplanes. A small number of statistics are then computed using the

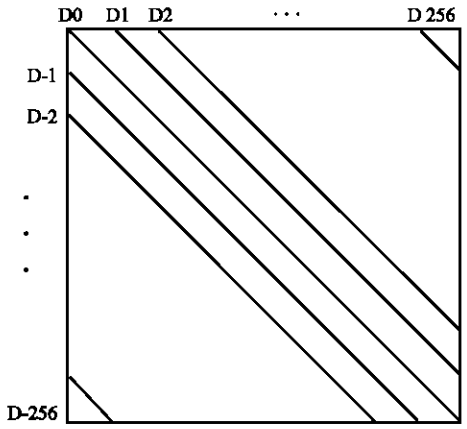

Fig. 8: Diagonals of co-occurrence matrix as feature

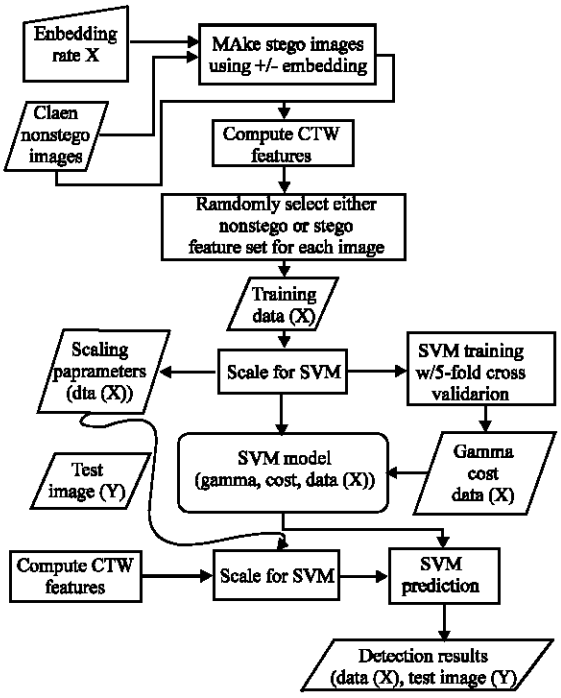

Fig. 9: Steganalysis process

model and fed into a support vector machine to classify detection results. Results presented are obtained using k-fold crossvalidation method using a large set of never compressed grayscale images.

Marvel *et al.* (2007) used lossless image compression to model the image and looks for discrepancies between the model for original images and for those containing steganography.

The entire process of feature computation, SVM training and testing/detection is shown in Fig. 9. The CTW features are extracted and the result is then scaled using the same scaling parameters specific to the

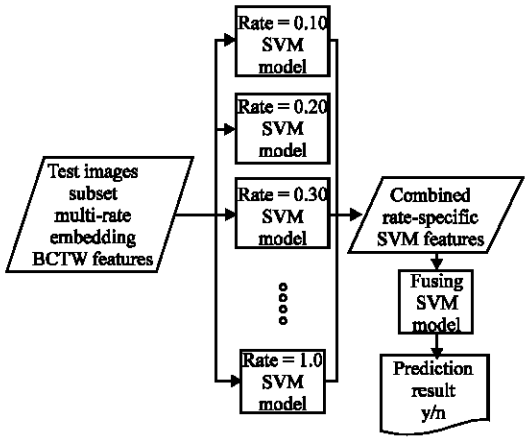

Fig. 10: Fusing classifiers

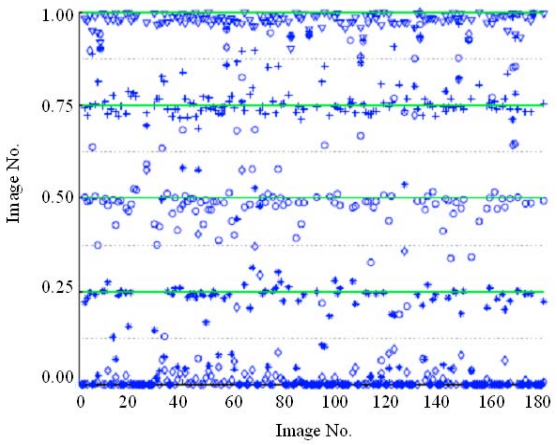

Fig. 11: Estimates of message lengths for ±1 embedding

model/classifier. These parameters are then input into the SVM prediction along with the model. The output of the detector is binary value representing a stego or non-stego prediction for each test image.

In the experimental work, a global detector that is trained using images with several steganographic embedding rates. Results show a small decrease in performance when employing the global detector. In most cases the performance of the global detector performs better than other embeddingrate mismatched detectors for the suspect images.

Marvel *et al.* (2008) further propose fusing multiple rate-specific SVMs in an attempt to improve upon the performance of the global classifier. SVM parameters from the rate-specific classifiers (e.g., distance from each models hyperplane) are used as input to the fusing classifier. A diagram for the fusing SVM is shown in Fig. 10.

The experiments show that both the global and fused rate-specific classifiers also work reasonably well, with the fused classifier performing somewhat better than the global classifier at higher embedding rates and at 50% true detection.

DEVELOPMENT OF ESTIMATORS FOR LSB MATCHING

Maximum likelihood estimator: Fridrich *et al.* (2005) presented a maximum likelihood estimator for estimating the number of embedding changes for non-adaptive ±K embedding in images. The method uses a high-pass FIR filter and then recovers an approximate message length using a Maximum Likelihood Estimator on those stego image segments where the filtered samples can be modeled using a stationary Generalized Gaussian random process. The results of detection are shown in Fig. 11.

It is shown that for images with a low noise level, such as decompressed JPEG images, this method can accurately estimate the number of embedding changes even for K = 1 and for embedding rates as low as 0.2 bits per pixel. Although, for raw, never compressed images the message length estimate is less accurate, when used as a scalar parameter for a classifier detecting the presence of ±K steganography, the proposed method gave us relatively reliable results for embedding rates as low as 0.5 bits per pixel. Unfortunately, the ML estimator starts to fail to reliably estimate the message length p once the variance of XF exceeds 9.

MAP estimator: At the same time, Holotyak *et al.* (2005a) proposed a new method for estimation of the number of embedding changes for non-adaptive ±k embedding in images. They present a stochastic approach based on sequential estimation of cover image and stego message. By modeling the cover image using the non-stationary Gaussian mode and the stego noise as additive mixture of random processes using Gaussian and Generalized Gaussian models. The stego message estimate is further analyzed using ML/MAP estimators to identify the pixels that were modified during embedding. For non-adaptive ±k embedding, the density of embedding changes is estimated from selected segments of the stego image. In Fig. 12, we show the block-diagram of the cover image estimation. The ML or Maximum A Posteriori (MAP) estimation were applied to estimate the parameters of the cover image model.

Experiments show that for images with a low level of noise (e.g., for decompressed JPEG images) this approach can detect and estimate the number of embedding changes even for small values of k, such as k = 2 and in some cases even for k = 1.

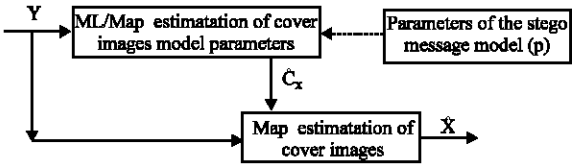

Fig. 12: Block diagram of cover image estimation method

CONCLUSIONS

In this study, we gave an overview of the detection methods for LSB matching steganography. As we can see, though some methods have been presented, the detection of LSB matching algorithm remains unresolved, especially for the uncompressed grayscale images. Moreover, new sophisticated steganographic methods will obviously require more refined detection methods. Steganalysis and steganography is just like a cat and mouse game and the steganalyzers will always be chasing the steganography developers. In the future, we will consider these challenging problems as an open field for future investigation as follows.

- Improving the detection performance for the case of low embedding ratio

When the embedding ratio is low, how to detect the existence of the secret message reliably is a difficult problem. Obviously, the detection accuracies of the existing methods are not enough, especially for the case of low embedding ratio.

- Improving the accuracy rate of estimator for LSB matching steganography

The existing estimating methods heavily relies on the fact that the embedding is non-adaptive and estimates the message length from those segments in the stego image that allow easier and more accurate modeling, such as flat or smooth areas. The Maximum Likelihood Estimator can accurately estimate the number of embedding changes for images with a low noise level, such as decompressed JPEG images. However, this approach is not effective for never-compressed images derived from a scanner. And the ML estimator "fail to reliably estimate the message length once the variance of the sample exceeds 9". Further improvement is expected by taking into consideration the cover image and the stego message stochastic models. It remains to be seen if these improvements will be sufficient for reliable and accurate estimation of secret message length in noisy images, such as never compressed images, scans, or certain resampled images.

- Looking for new methods of image feature extraction

Extract more informative features to detect the existence of secret messages embedded with most kinds of steganography methods. Although a number of features have been found out, they are not effective enough to have desirable accuracy for most embedding schemes.

- Improving the detection performance of blind steganalysis

Nowadays, image blind steganalysis is still challenging in many aspects. And the existing blind steganalysis are far from being applied in reality.

- Identifying the image modified by steganography or normally processing operation

Usually some normal image processing operations, such as images splicing, stretching, smoothing, sharpening, erosion, dilation and so on, always destroy the statistical characteristics of natural images and lead to the wrong detection. How to distinguish the image modified by normal image processing operation or steganography is a new challenge for steganalyzers.

ACKNOWLEDGMENTS

This project is supported by Scientific Research Fund of Hunan Provincial Education Department (Grant No. 09B019), Hunan Provincial Natural Science Foundation of China (Grant No. 09JJ4033), Science and Technology Program of Hunan Province (Grant No. 2010FJ3090) and Science and Technology Program of Yiyang City (Grant No. YK0956).

REFERENCES

Abolghasemi, M., H. Aghainia, K. Faez and M.A. Mehrabi, 2008. Steganalysis of LSB matching based on co-occurrence matrix and removing most significant bit planes. Proceedings of the 2008 International Conference on Intelligent Information Hiding and Multimedia Signal Processing, Aug. 15-17, IEEE Computer Society, Washington, DC., USA., pp: 1527-1530.

Avcibas, I., N. Memon and B. Sankur, 2003. Steganalysis using image quality metrics. IEEE Trans. Image Process, 12: 221-229.

Boncelet, C. and L. Marvel, 2007. Steganalysis of ±1 embedding using lossless image compression. Proceedings of the IEEE International Conference on Image Processing, Sept. 16, Delaware University, Newark, pp: 149-152.

Cancelli, G., M. Barni, G. Doerr and I.J. Cox, 2008. A comparative study of ±1 steganalyzers. Proceedings of the IEEE International Workshop on Multimedia Signal Processing, Jan. 27-31, Queensland, Australia, pp: 791-796.

Chen, C. and Y.Q. Shi, 2008. JPEG image steganalysis utilizing both intrablock and interblock correlations. Proceedings of the IEEE International Symposium on Circuits and Systems, May 18-21, IEEE Computer Society Press, Washington, USA., pp: 3029-3032.

Dumitrescu, S., X.L. Wu and Z. Wang, 2003. Detection of LSB steganography via sample pair analysis. IEEE Trans. Signal Process., 51: 1995-2007.

Farid, H., 2002. Detecting hidden messages using higher-order statistical models. Proceedings of the IEEE International Conference on Image Processing, Sept. 22-25, IEEE Computer Society Press, New York, USA., pp: 905-908.

Farid, H. and S. Lyu, 2003. Detecting hidden messages using higher-order statistics and support vector machines. Proceedings of the 5th International Information Hiding Workshop LNCS, Berlin, Oct. 7-9, Springer-Verlag, pp: 340-354.

Fridrich, J., R. Du and M. Long, 2000. Steganalysis of LSB encoding in color images. Proceedings of the IEEE International Conference on Multimedia and Expo, July 31-Aug. 2, USA., New York, pp: 1279-1282.

Fridrich, J., M. Goljan and R. Du, 2001. Detecting LSB steganography in color and gray-scale images. Multimedia IEEE, 8: 22-28.

Fridrich, J., D. Soukal and M. Goljan, 2005. Maximum likelihood estimation of length of secret message embedded using ±K steganography in spatial domain. Proc. SPIE, 5681: 595-606.

Goljan, M., J. Fridrich and T. Holotyak, 2006. New blind steganalysis and its implications. Proc. SPIE, 6072: 607201-607201.

Harmsen, J. and W. Pearlman, 2003. Steganalysis of additive noise modelable information hiding. Proc. SPIE, 5022: 131-142.

Harmsen, J.J., K.D. Bowers and W.A. Pearlman, 2004. Fast additive noise steganalysis. Proc. SPIE., 5306: 489-495.

Holotyak, T., J. Fridrich and D. Soukal, 2005a. Stochastic approach to secret message length estimation in ±k embedding steganography. Proc. SPIE., 5681: 673-684.

Holotyak, T., J. Fridrich and S. Voloshynovskiy, 2005b. Blind statistical steganalysis of additive steganography using wavelet higher order statistics. Proceedings of the 9th IFIP TC-6 TC-11 Conference on Communications and Multimedia Security, Sept. 19-21, Salzburg, Austria, pp: 273-274.

Hu, L.N., L.G. Jiang and C. He, 2008. A novel steganalysis of lsb matching based on kernel fda in grayscale images. Proceedings of the IEEE International Conference Neural Networks and Signal Processing, (IEEEICNSP'08), IEEE, Zhenjiang, China, pp: 556-559.

Huang, F.J., B. Li and J.W. Huang, 2007. Attack LSB matching steganography by counting alteration rate of the number of neighbourhood gray levels. Proceedings of the IEEE International Conference on Image Processing, Oct. 16-19, San Antonio, pp: 401-404.

Jiao-Hua, Q., S. Xing-Ming, C. Xiao-Yan, 2007a. Steganalysis based on lifting wavelet transform for palette images. Proceedings of the International Conference on Computational Intelligence and Security Workshops, Dec. 15-19, IEEE Xplore London, pp: 672-675.

Jiao-Hua, Q., S. Xing-Ming, C. Xiao-Yan, 2007b. Steganalysis based on statistical characteristic of adjacent pixels for LSB steganography. J. Syst. Simulat., 19: 5856-5860.

Ker, A., 2004a. Improved detection of LSB steganography in grayscale images. Lecture Notes Comput. Sci., 3200: 97-115.

Ker, A., 2004b. Quantitative evaluation of pairs and RS steganalysis. Proc. SPIE, 5306: 83-97.

Ker, A., 2005a. Steganalysis of LSB matching in grayscale images. IEEE Signal Process. Lett., 12: 441-444.

Ker, A.D., 2005b. Resampling and the detection of LSB matching in colour bitmaps. Proc. SPIE., 5681: 1-15.

Liu, Q., A.H. Sung and B. Ribeiro, 2005. Statistical correlations and machine learning for steganalysis. Proceedings of the Conference on Adaptive and Natural Computing Algorithms, (CANCA'05), Springer, Wien, New York, pp: 437-440.

Liu, Q., A.H. Sung, J. Xu and B.M. Ribeiro, 2006. Image complexity and feature extraction for steganalysis of LSB matching steganography. Proc. 18th Int. Conf. Pattern Recognition, 2: 267-270.

Liu, Q., A.H. Sung, B. Ribeiro, M. Wei, Z. Chen and J. Xu, 2008. Image complexity and feature mining for steganalysis of least significant bit matching steganography. Inform. Sci., 178: 21-36.

Luo, X.Y., D.S. Wang, P. Wang and F.L. Liu, 2008. A review on blind detection for image steganography. Signal Process., 88: 2138-2157.

Lyu, S. and H. Farid, 2004. Steganalysis using color wavelet statistics and oneclass vector support machines. Proc. SPIE, 5306: 35-45.

Marvel, L., B. Henz and C. Boncelet, 2007. A performance study of ±1 steganalysis employing a realistic operating scenario. Proceedings of the IEEE Military Communications Conference MILCOM, Oct. 29-31, Orlando, FL., USA., pp: 1-7.

Marvel, L., B. Henz and C. Boncelet, 2008. Fusing rate-specific SVM classifiers for ±1 embedding steganalysis. Proceedings of the 42th Annual Conference on Information Sciences and Systems, March 19-21, Princeton, pp: 361-364.

Miche, Y., B. Roue, A. Lendasse and P. Bas, 2006. A feature selection methodology for steganalysis. Proceedings of the International Workshop on Multimedia Content Representation, Classification and Security, LNCS 4507, Sept. 11-13, Springer-Verlag, Istanbul, Turkey, pp: 49-56.

Mielikainen, J., 2006. LSB matching revisited. IEEE Signal Process. Lett., 13: 285-287.

Niu, C.M., X.M. Sun, J.H. Qin and Z.H. Xia, 2009. Steganalysis of two least significant bits embedding based on least square method. Proceedings of International Conference on Computing, Communication, Control and Management, Aug. 8-9, Sanya, China, pp:124-127.

Pevni, T. and J. Fridrich, 2007. Merging markov and DCT features for multi-class JPEG steganalysis. Proceedings of the SPIE, Electronic Imaging, Security, Steganography and Watermarking of Multimedia Contents IX, Jan. 29, SPIE, San Jose, California, pp: 301-313.

Qin, J., X. Sun, X. Xiang and Z. Xia, 2009a. Steganalysis based on difference statistics for LSB matching steganography. Inform. Technol. J., 8: 1281-1286.

Qin, J., X. Sun, X. Xiang, C. Niu, 2009b. A principal feature selection and fusion method for image steganalysis. J. Elect. Imag., 18: 1-14.

Shi, Y.Q., C. Chen and W. Chen, 2006. A markov process based approach to effective attacking JPEG steganography. Proceedings of the 8th Information Hiding Workshop, July 10-12, Springer-Verlag, Brittany, France, pp: 249-264.

Wang, Y. and P. Moulin, 2007. Optimized feature extraction for learning-based image steganalysis. IEEE Trans. Inform. Forensics Security, 2: 31-45.

Westfeld, A., 2002. Detecting low embedding rates. Proceedings of the Information Hiding Workshop, Oct. 07-09, Springer-Verlag, London, UK., pp. 324-339.

Xia, B., X.M. Sun and J.H. Qin, 2009a. Steganalysis based on neighbourhood node degree histogram for LSB matching steganography. Proceedings of the 1st International Conference on Multimedia Information Networking and Security, Nov. 18-20, Wuhan, China, pp: 79-82.

Xia, Z., X. Sun, J. Qin, C. Niu, 2009b. Feature selection for image steganalysis using hybrid genetic algorithm. Inform. Technol. J., 8: 811-820.

Xia, Z.H., X.M. Sun, W. Liang, J.H. Qin and F. Li, 2010. JPEG image steganalysis using joint DCT domain features. J. Electronic Imaging, 19: 1-13.

Xuan, G.R., Y.Q. Shi, J.J. Gao, D.K. Zou and C.Y. Yang *et al*., 2005. Steganalysis based on multiple features formed by statistical moments of wavelet characteristic functions. Proceedings of the 7th International Information Hiding Workshop, LNCS 3727, June 6-8, Springer-Verlag, Berlin, pp: 262-277.

Yu, X.Y. and N. Babaguchi, 2008a. An improved steganalysis method of LSB matching. Proceedings of the Intelligent Information Hiding and Multimedia Signal Processing, Aug. 15-17, Osaka University, Suita, pp: 557-560.

Yu, X.Y. and N. Babaguchi, 2008b. Run length based steganalysis for LSB matching steganography. Proceedings of the IEEE International Conference on Multimedia and Expo, June 23-April 26, Hannover, Germany, pp: 353-356.

Zhang, J., I.J. Cox and G. Doërr, 2007. Steganalysis for LSB matching in images with high-frequency noise. Proceedings of the IEEE 9th Workshop on Multimedia Signal Processing, Oct. 1-3, Piscataway, New Jersey, USA., pp: 385-392.