ATTACK LSB MATCHING STEGANOGRAPHY BY COUNTING ALTERATION RATE OF THE NUMBER OF NEIGHBOURHOOD GRAY LEVELS

Fangjun Huang[1,2], Bin Li[1,2], Jiwu Huang[1,2]

1. Guangdong Key Lab. of Information Security
2. Dept. of Electronics and Communication Engineering,
Sun Yat-Sen University, Guangzhou, Guangdong, 510275, P. R. China

ABSTRACT

In this paper, we propose a new method for attacking the LSB (least significant bit) matching based steganography. Different from the LSB substitution, the least two or more significant bit-planes of the cover image would be changed during the embedding in LSB matching steganography and thus the pairs of values do not exist in stego image. In our proposed method, we get an image by combining the least two significant bit-planes and divide it into 3×3 overlapped subimages. The subimages are grouped into four types, i.e. T_1, T_2, T_3 and T_4 according to the count of gray levels. Via embedding a random sequence by LSB matching and then computing the alteration rate of the number of elements in T_1, we find that normally the alteration rate is higher in cover image than in the corresponding stego image. This new finding is used as the discrimination rule in our method. Experimental results demonstrate that the proposed algorithm is efficient to detect the LSB matching stegonagraphy on uncompressed gray scale images.

Index Terms—Steganography, Steganalysis, LSB matching, LSB substitution

1. INTRODUCTION

Image steganography concerns the problem of hiding information in a carrier. In spatial domain, LSB-based steganography is a conventional method and can be divided into two classes, the LSB substitution [1] and LSB matching [2].

LSB substitution is a method that the LSBs of the cover pixels are simply overwritten by the secret bit stream. According to [3], LSB substitution modifies only the least significant bits and generates pairs of values (PoVs) in the stego image. Since the secret message is generally encrypted first for security and thus uniformly distributed, we can suppose that in the stego image the frequencies of both values of each PoV are equal if the cover image is fully embedded with secret message. Some reported methods, such as Chi-square attack [3], RS analysis [4] and Sample Paris analysis [5] can not only detect stego images with high reliability, but also estimate the length of secret message accurately.

However, the LSB matching, a counterpart of LSB substitution, makes the detection of hidden message much more difficult. The LSB matching does not simply overwrite the LSBs. Instead, the randomly selected sample value is increased or decreased if its LSB does not match the secret message bit to be embedded. The LSB of the cover pixel value finally equals the next bit of the hidden data and PoVs are not equalized. In [5], Sharp described that the effect of LSB matching on the histogram of the sample values is identical to filtering it with a low-pass filter {0.25, 0.5, 0.25}. Without access to the cover image, it is difficult to recognize this filtering effect, especially when the hidden bit rate is low. The Chi-square attack, RS analysis and Sample Pairs analysis will do not work as described in [6, 8, 9]. Recently, a few of detecting methods have been reported. In [6], Westfeld presented a detector for color images. It considers the count of neighbor colors, since a color in a compressed cover image has only 4 or 5 neighbors on average and in the stego image achieved by LSB matching may have more than 9 neighbors. However, the proposed algorithm showed poor performance for the uncompressed color images and gray scale images. Another detector applicable to LSB matching was presented by Harmsen and Pearlman [7], in which the embedding process was modeled as the addition of noise. Under an independence assumption, the histogram of the stego image is a convolution of the noise probability mass function (PMF) and the original histogram. Harmsen and Pearlman proved that the embedding methods are quantified by a decrease in the center of mass (COM) of the histogram characteristic function (HCF) in the frequency domain. Ker [8] extended Harmsen's method on LSB matching and concluded that such a detector performed poorly indeed,

This work was supported by NSFC (60325208, 90604008, 60633030), 973 Program (2006CB303104), NSF of Guangdong (04205407), and Young teacher's fund of Sun Yat-Sen University (1131159).
*Correspondence author: huangfj@mail.sysu.edu.cn

especially for gray scale images. 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. However, the experiments demonstrated that it was not very efficient. Fridrich, et.al [9, 10] presented some new methods for attacking the $\pm K$ steganography, which is efficient in detecting the LSB matching using decompressed JPEGs. However, as described in [9], for some raw, never compressed images from a scanner, the embedded message in any of the images were unable to detect. As we can see, though a lot of methods have been presented, the detection of LSB matching algorithm remains unresolved, especially for the uncompressed grayscale images.

In this paper, a new method for attacking the LSB matching steganography using uncompressed gray scale images is presented. For a given image, we 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 T_1 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 T_1, we 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 our method. Experimental results demonstrate that the proposed method is efficient to detect the LSB matching stegonagraphy using uncompressed gray scale images.

The rest of this paper is organized as follows. Section 2 describes the proposed method. Section 3 shows our experiments. The paper is concluded in Section 4.

2. THE PROPOSED DETECTION ALGORITHM

According to some steganographic experts' opinion [3, 11], the uncompressed grayscale image is the most secure cover image for steganography. In our works, we only consider the uncompressed grayscale images. Suppose an $M \times N$ grayscale image $I(x,y)$ is composed of eight 1-bit planes $I_0 \sim I_7$, ranging from bit-plane 0 for the least significant bit to bit-plane 7 for the most significant bit. Since the LSB matching method mainly influence the least two significant bit-planes, we get an image $A(x,y)$ by combining the least two significant bit-planes as follows.

$$A(x,y) = I_0(x,y) + I_1(x,y) \times 2 \qquad (1 \le x \le M, 1 \le y \le N)$$

There are only four gray levels in the image, i.e. 0, 1, 2 and 3.

In our proposed method, we use a 3×3 subimage to define the neighborhood of a point $A_{i,j}(2 \le i \le M-1, 2 \le j \le N-1)$ in $A(x,y)$. The center of the subimage is moved pixel by pixel starting at the top left corner. At each position, we count the number of gray levels in this 3×3 subimage. There are four types of subimages.

T_1 : Including only one gray level. That is, all the pixels in the subimage belonging to this type have the same value.

T_2 : Including two gray levels. That is, all the subimages in this type have two gray levels.

T_3 : Including three gray levels. That is, all the subimages in this type have three gray levels.

T_4 : Including four gray levels. That is, all the subimages in this type have four gray levels.

We propose to classify the cover and stego image using the following idea. In the embedding procedure, the probability that the subimages changing from T_1 to $T_i(2 \le i \le 4)$ is much more than the probability that the subimages changing from $T_i(2 \le i \le 4)$ to T_1. For example, given a subimage belonging to T_1, if any one pixel is shifted, the probability that the subimage would belong to $T_i(2 \le i \le 4)$ is 100%. However, for any subimage belonging to $T_i(2 \le i \le 4)$ initially, if one pixel is modified, the probability that the subimage would belong to T_1 is much less than 100%. Thus we conclude that $|T_1|$ would decrease after the embedding, where $|T_1|$ denotes the number of elements belonging to T_1.

However, it is not an easy work to define a threshold for the number $|T_1|$. Due to the different kinds of images, the values of $|T_1|$ distribute in a very wide range. Fortunately, we have made an important observation that enables us to reliably distinguish the cover and stego images. Suppose $|T_1^C|$ is the number of elements belonging to T_1 of the cover image $I_C(x,y)$ and $|T_1^S|$ is the number of elements belonging to T_1 of the corresponding stego image $I_S(x,y)$. Embedding another random sequence into the cover and stego images by LSB matching simultaneously, we can obtain $|T_1^{C*}|$ and $|T_1^{S*}|$, which are the count of elements belonging to T_1 in the obtained images. Denote the alteration rate as

$$k_c = \frac{\left|T_1^C\right| - \left|T_1^{C*}\right|}{\left|T_1^C\right|}$$

and

$$k_s = \frac{\left|T_1^S\right| - \left|T_1^{S*}\right|}{\left|T_1^S\right|}$$

It is noted that

$$k_c > k_s \quad (1)$$

usually holds true if the stego image contain a large amount of hidden data. Some explanation can be found in Fig. 1. There are nine 3×3 overlapped subimages, where $A_1 \sim A_9$ are the centered pixels of the corresponding subimages. For simplicity, only three subimages B_3, B_5 and B_7 are labeled. According to our various experiments, the subimages belonging to T_1 are concentrated in the regions with little texture complexity. Firstly we assume Fig. 1 is a sub region of $A(x, y)$ obtained from a cover image and all the nine subimages belong to T_1. Since the subimage is overlapped, two or more sumimages will change from T_1 to $T_i (2 \le i \le 4)$ if we embed one random bit into any of the pixels except the four pixels, i.e. A_{LT}, A_{LB}, A_{RT} and A_{RB}. Especially, when we embedded a random bit into the pixel A_5, all the nine sumimages will change from T_1 to $T_i (2 \le i \le 4)$. However, if Fig. 1 is a sub region obtained from a stego image and some secret message bits have been embedded. The subimages belonging to T_1 would not be overlapped as before and some subimages would belong to $T_i (2 \le i \le 4)$. When one random bit is embedded, the number of subimages changed from T_1 to $T_i (2 \le i \le 4)$ will decrease greatly. From the above analyses, we can deduce that Eq. (1) usually holds true. 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 as described in Fig. 1, it is difficult for us to distinguish the cover and stego images using Eq. (1) as a discrimination rule.

Based on above observation, we present the following steps for detection algorithm.

(1) For any given image, calculate $\left|T_1\right|$.

(2) Embed a random sequence with a length *l* into the given image by LSB matching. For example, *l*=0.5 means for every two pixels of the given image, one message bit is embedded.

(3) Calculate $\left|T_1^*\right|$ of this new obtained image.

(4) We get the alteration rate *k* by using

$$k = \frac{\left|T_1\right| - \left|T_1^*\right|}{\left|T_1\right|}$$

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

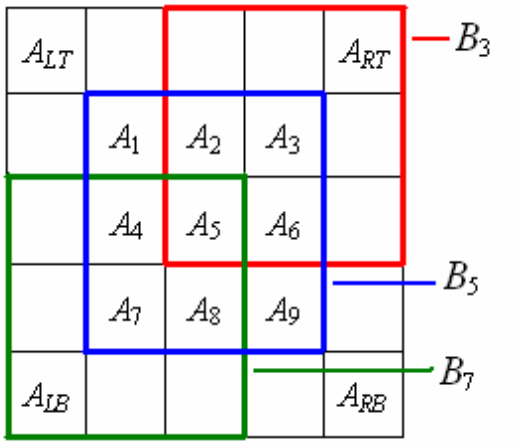

Fig. 1 The demonstration of overlapped subimages

3. EXPERIMENTAL RESULTS

Experimental results are given in this section to demonstrate the performance of our proposed method. We test our algorithm on 1338 images from an uncompressed image database, UCID [12]. All the images in UCID are high resolution TIFF files with the size 512×384 or 384×512 and appear to be cut from a variety of uncompressed digital camera images. Before testing, the images were converted to grayscale using the rgb2gray function of Matlab 7.0 directly. The random sequence that we use to get the alteration rate *k* is chosen with a length *l*=0.3.

In our experiments, we find that the values of $\left|T_1\right|$ mainly distribute in the region [10, 10^4]. For about 0.2% cover and more than 0.7% stego images (with the secret message length *p*=0.25, 0.5, 0.75, 1), $\left|T_1^*\right|$ becomes zero after the random sequence is embedded. All these images are discriminated as stego images in our proposed method. Fig. 2 shows the receiver operation characteristic (ROC), where the horizontal axis represents the probability of false positive and vertical axis represents the probability of detection. The results corresponding to different stego images (the secret message length *p*=0.25, 0.5, 0.75, 1 respectively) are given. We also test our scheme on the UCID database with the image size 640×480 or 480×640 [13], the results is similar. Since the alteration rate is a

dimensionless discriminator, our scheme can be used to detect images with different size. Moreover, for a given detection rate, the threshold is relatively stable for images with different size.

We compare our method with Ker's two methods presented in [8]. As described in [8], the calibration technique is efficient only when the secret message is fully embedded, thus the stego images with the secret message length p=1 are adopted in our comparison. The experimental results are shown in Fig. 3. As we can see, under the same probability of false positive, the detection rate of our method is much higher than Ker's two methods. Moreover, we can combine our technique with Ker's to achieve a more reliable detection rate. We will refer it in our future work.

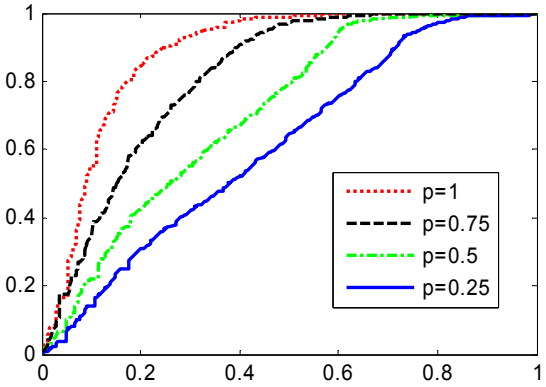

Fig. 2 The detection results corresponding to different message length

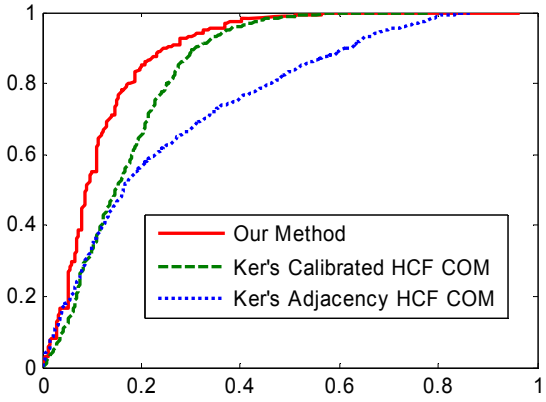

Fig. 3 The results compared with Ker's method

4. CONCLUSIONS

In this paper, we have proposed a new method for detection of LSB matching steganography, and tested its performance on a database of uncompressed gray scale images. The main merits of our method are as follows.

(1) According to the properties of LSB matching, we find a novel discrimination rule to distinguish the cover and stego images, in which only the least two significant bit-planes need to be considered.

(2) Since the alteration rate is a dimensionless discriminator, we can use our scheme to detect the given images with different size.

(3) It is easy to combine our technique with Harmsen's or Ker's to design a more reliable detection rule.

5. REFERENCES

[1] F. A. P. Petitcolas, R. J. Anderson and M. G. Kuhn, "Information hiding-a survey", *Proc.of IEEE*, vol. 87, no. 7, pp. 1062-1078, 1999.

[2] T. Sharp, "An implementation of key-based digital signal steganography", *Proc. 4th International workshop on Information Hiding*, LNCS 2137, pp. 13-26, Springer-Verlag, 2001

[3] A. Westfeld and A. Piftzmann, "Attacks on Steganographic Systems", *Proc. 3rd International Workshop on Information Hiding*, *LNCS 1768*, pp. 61-76, Springer-Verlag, 2000

[4] 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, 2001

[5] S. Dumitrescu, X. Wu, and Z. Wang. "Detection of LSB steganography via sample pair analysis", *IEEE Trans. Signal Processing*, vol. 51, no. 7, pp. 1995-2007, 2003

[6] A. Westfeld, "Detecting low embedding rates", *Proc. 5th International Workshop on Information Hiding*, LNCS 2578, pp. 324-339, Springer-Verlag, 2003

[7] J. J. Harmsen and W. A. Pearlman, "Steganalysis of additive noise modelable information hiding", *Proc. SPIE, Security, Steganography, and Watermarking of Multimedia Contents* V, vol. 5020, pp. 131-142, 2003

[8] A. D. Ker, "Steganalysis of LSB matching in grayscale images", *IEEE Signal processing letters*, vol. 12, no. 6, pp. 441-444, 2005

[9] J. Fridrich, D. Soukal and M. Goljan, "Maximum likelihood estimation of length of secret message embedded using $\pm K$ steganography in spatial domain", *Proc. SPIE, Security, Steganography, and Watermarking of Multimedia Contents* VII, vol. 5681, pp. 595-606, 2005

[10] T. Holotyak, J. Fridrich and David Soukal, "Stochastic approach to secret message length estimation in $\pm k$ embedding steganography", *Proc. SPIE, Security, Steganography, and Watermarking of Multimedia contents* VII, vol. 5681, pp. 673-684, 2005

[11] T. Aura, "Practical invisibility in digital communication", *Proc. 1rd International Workshop on Information Hiding*, LNCS 1174, pp. 265-278, Springer-Verlag, 1996.

[12] G. Schaefer and M. Stich, "UCID - An Uncompressed Colour Image Database", *Proc. SPIE, Storage and Retrieval Methods and Applications for Multimedia*, vol. 5307, pp. 472-480, 2004

[13] G. Schaefer and M. Stich, UCID - An Uncompressed Colour Image Database, *Technical Report, School of Computing and Mathematics*, Nottingham Trent University, U.K, 2003.