Mathematical Biosciences and Engineering

http://www.aimspress.com/journal/MBE

MBE, 16(5): 4908–4922
DOI: 10.3934/mbe.2019247
Received: 26 January 2019
Accepted: 25 April 2019
Published: 30 May 2019

Research article

Locating secret messages based on quantitative steganalysis

Chunfang Yang[1], Fenlin Liu[1], Shuangkui Ge[2], Jicang Lu[1,*]and Junwei Huang[3]

[1] State Key Laboratory of Mathematical Engineering and Advanced Computing, Zhengzhou, Henan, 450001, China

[2] Beijing Institute of Electronic Technology Application, Beijing, 100000, China

[3] HERE North American LLC, Burlington, Massachusetts, 01803, USA

* **Correspondence:** Email: lujicang@sina.com.

Abstract: Steganography poses a serious challenge to forensics because investigators cannot identify even traces of secret messages embedded using a steganographer. Contrarily, the objective of locating steganalysis is to locate the embedded message, which should help extract the secret message. In this paper, a methodology of locating steganalysis using quantitative steganalysis is presented for multiple stego images with embedded messages along the same embedding path. Three typical quantitative steganalysis methods are applied to the methodology to locate the messages embedded using LSB replacement. Experimental results show that the presented methods can reliably estimate the embedding positions, which verifies the validity of the presented methodology. The presented methodology points out a new use of quantitative steganalysis, and further demonstrates that it is necessary to design more precise quantitative steganalysis methods.

Keywords: steganalysis; steganography; locating steganalysis; embedding path

1. Introduction

Steganography is the art of hiding the very presence of communication by embedding secret messages into innocuous looking cover objects, such as digital images, audios, videos, and text files [1, 2, 3, 4]. A large number of tools have been developed based on steganography techniques, and some of them can be downloaded from the Internet for free and used without advanced knowledge of steganography. This poses a serious challenge to forensics and security.

Consequently, steganalysis has attracted a lot of attention as its goal is to detect stego signals [5, 6]. In theory, steganography is considered broken when the presence of a secret message can be established [1]. In real-world scenarios, steganalysts are usually interested in more than the detection of a secret message, and their ultimate goal is to extract the secret message. Therefore, Fridrich et al. proposed the procedure of steganalysis on a more general level in [7]. To extract the secret message, steganalysts still need more details, such as the length of secret message, modification ratio of the cover signal [7, 8], or positions where the message bits are embedded.

Steganalysis that can estimate the length of a secret message or the modification ratio of signal samples is called quantitative steganalysis [9]. In 2003, Fridrich et al. [9] proposed a principle which can be used to develop quantitative steganalysis methods. Currently, numerous quantitative steganalysis methods have been proposed, such as structural steganalysis [10, 11, 12, 13, 14, 15] and weighted stego image (WS) steganalysis [16]. However, these quantitative steganalysis methods were only designed for some popular steganography methods, and they treat each steganography method individually. Therefore, in 2009, Pevny et al. [17] presented a general methodology for constructing quantitative steganalyzers from features used for blind steganalysis. The presented methodology has attracted much attention because it is otherwise difficult to analyze the relationship between statistical features and the embedding ratio for many steganography methods [18].

Steganalysis that can locate embedded messages is called locating steganalysis. If steganalysts can locate the embedded messages using locating steganalysis, then the extraction of these messages should be converted to a cryptanalysis problem. Researchers have proposed some methods to utilize the results of locating steganalysis to estimate the groups of group parity steganography [19], determine the stego pixels order of random steganography [20], and even recover the stego key [21]. Therefore, locating steganalysis will play an important role in the forensics of steganography. When compared with quantitative steganalysis, there is little literature on locating steganalysis because of its difficulty.

The early research on locating steganalysis mainly focuses on estimating the embedding positions of sequential steganography. In 1999, Westfeld et al. [22] presented the χ^2 method for sequential least significant bit (LSB) replacement. In 2005, Trivedi et al. [23] presented the cumulative sum test method and applied it to the spread spectrum steganography. In 2007, Ker [24] applied weighted stego image steganalysis to estimating the embedding positions of sequential LSB replacement; etc.

For random steganography, the limited existing literatures are aimed at some special cases. Fridrich et al. [7] assumed that steganalysts already know (or suspect) the steganographic algorithm under investigation and the Pseudo-Random Number Generator (PRNG) that is seeded with a seed derived from a user-specified stego key to generate a pseudo-random path. They then presented a search methodology to determine not only the embedding positions, but also the correct stego key by quantifying statistical properties of samples along portions of the embedding path. The presented methodology was applied to JPEG steganography and spatial steganography [7, 8]. Ker et al. [25, 26] assumed that steganalysts possess a number of stego images, each containing message at the same locations with LSB replacement or LSB matching, then utilized the weighted stego image residuals or wavelet absolute moments (WAM) residuals to locate the messages respectively. Luo et al. [27] improved the WS residuals and proposed two novel residual calculation methods to locate the stego positions of LSB matching with higher accuracy than the algorithm in [26]. Quach [28, 29] modelled the image using the Hidden Markov model, adopted the maximum a posteriori (MAP) method to estimate the cover image, then located the stego position of LSB steganography by combining the differences between multiple stego images and their corresponding estimated cover images. Gui et al. [30] estimated 9 cover images by averaging 4-neighborhood pixels and adopting MAP estimation along eight directions, then calculated 9 residuals for every pixel to locate the payload of LSB matching and achieved higher performance than the algorithm proposed in [28, 29]. Later, Quach [31] proposed an algorithm based on Markov Random Fields to estimate the cover image by considering the inherent two-dimensional nature of images, which improves the locating accuracy for LSB steganography. Liu et al. [32] proposed an algorithm to estimate the cover image by compressing the stego image which has suffered JPEG compression before embedding the message into the LSBs, which can be used to locate the payload with higher accuracy. Yang et al. [33] proved the optimal stego subset property of multiple least significant bits (MLSB) steganography, and they proposed two algorithms based on this property to extract the hidden message by recovering the stego key and locating the payload for two different cases respectively.

In conclusion, there are a number of locating steganalysis methods for typical steganography methods, some of which have been used to extract the embedded message under strict assumptions, such as sequential steganography, and knowing the steganography algorithm and the PRNG. However, there is still no universal methodology which can guide the design of a concrete locating steganalysis algorithm for a given steganography algorithm. Considering the idea from blind to quantitative steganalysis [17], it should be possible to apply the results of quantitative steganalysis to locating steganalysis. Therefore, this paper focuses on research on "from quantitative to locating steganalysis". The main contributions of this paper are as follows:

1) A methodology of locating steganalysis is proposed using quantitative steganalysis under the condition of owning multiple stego images embedded along the same embedding path.

2) The typical structural steganalysis and weighted stego image steganalysis algorithms are applied to the proposed methodology. Then, three concrete locating steganalysis algorithms are designed for LSB replacement steganography.

3) The designed locating steganalysis algorithms are tested experimentally. The experimental results show that the designed locating steganalysis algorithms can estimate the embedding positions reliably. Specifically, the locator from sample pair analysis (SPA) can locate more than 90% of the stego positions when the number of stego images owned is more than 500.

The remainder of this paper is structured as follows: Section 2 will describe the principle of how to use quantitative steganalysis to locate the stego positions, and propose the methodology of locating steganalysis from quantitative steganalysis. Section 3 will apply two categories of typical quantitative steganalysis methods to the proposed methodology, and design some locating steganalysis algorithms for LSB replacement. Section 4 will supply the experimental results and analysis for the designed locating steganalysis algorithms. Finally, the paper is summarized in Section 5.

2. Methodology of locating steganalysis from quantitative steganalysis

In actual communication, the participants usually keep the same stego key during a time interval. Then, multiple stego objects with the same stego key may be obtained. For these stego objects, the same stego key can generate the same embedding paths for some cases, for example,

1) The secret messages are embedded in sequence with the start position as stego key.

2) The pixels in a cover image are permuted by the stego key, then the secret message is embedded into the permuted pixels in sequence. At this time, the same stego key and number of samples in stego objects will generate the same embedding paths.

Under the condition of owning multiple stego images embedded along the same embedding path, this section analyzes the difference between embedding ratios in stego position and non-stego position. Then, a method is designed to estimate the embedding ratio in each position by using quantitative steganalysis. Finally, a methodology of locating steganalysis from quantitative steganalysis is proposed. In the remainder of this paper, the following symbols will be used:

T: the number of stego images embedded along the same path;

n: the number of pixels in each stego image;

b: the number of bits used to store a pixel;

S_t=$\{s_{t,1}, s_{t,2}, \ldots, s_{t,n}\}$: the t-th stego image which is scanned row by row, $1 \leq t \leq T$;

$s_{t,i}$: the i-th pixel in the t-th stego image, $0 \leq s_{t,i} \leq 2^b - 1$, $1 \leq t \leq T$, $1 \leq i \leq n$;

$L(t)$: the length of the secret message in the t-th stego image, $1 \leq t \leq T$;

l_{min}: the minimum length of a secret message in the T stego images;

l_{max}: the maximum length of a secret message in the T stego images;

$Path(key) = \langle K(1),\ K(2),\ \ldots,\ K(n)\rangle$: the embedding path generated by stego key key;

$K(j)$: the position of the pixel where the j-th bit of the secret message is embedded, $1 \leq j \leq n$, $1 \leq K(j) \leq n$, and $K(j) \neq K(k)$ for $j \neq k$;

$f(i)$: the ratio of the number of stego pixels to the i-th pixels of all T stego images.

2.1. Analysis of embedding ratio in each position

When embedding T secret messages into T images along the same embedding path, two different cases may occur as shown in Figure 1 and Figure 2.

a) **The lengths of T secret messages are all equal to l, viz. $l_{min} = l_{max} = l$.** Then it can be seen from Figure 1 that if the position i is one of the first l positions in the embedding path, viz. there is $K(j)(j \leq l)$ which is equal to i, then all of the i-th pixels of T stego images contain secret messages and the embedding ratio of the i-th pixels of T stego images is equal to 1, viz. $f(i) = 1$; otherwise, all of the i-th pixels of T stego images do not contain secret messages, and the embedding ratio of the i-th pixels of T stego images is equal to 0, viz. $f(i) = 0$.

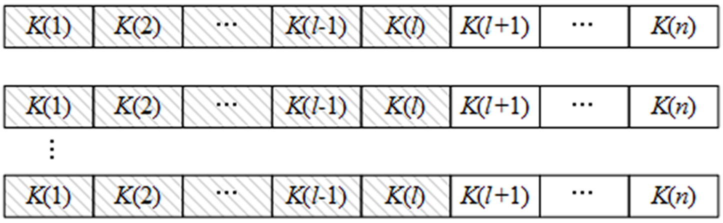


Figure 1. Embedding paths of T stego images embedded secret messages of equal length l, where the squares with oblique lines denote the stego positions.

b) **The lengths of T secret messages are different, and the minimum length of these secret messages is equal to l, viz. $l_{min} = l$.** Then it can be seen from Figure 2 that if the position i is one of the first l positions in the embedding path, all of the i-th pixels of T stego images contain secret messages and the embedding ratio of the i-th pixels of T stego images is equal to 1, viz. $f(i) = 1$; otherwise, the i-th pixels of T stego images do not all contain secret messages and the embedding ratio of the i-th pixels of T stego images is smaller than 1, viz. $f(i) < 1$. Furthermore, when the position i in the embedding path is backward in order, $f(i)$ will decrease gradually until it is equal to 0, viz. when $l < j < k \leq l_{\max}$, $1 > f(K(j)) \geq f(K(k)) > f(K(l_{\max} + 1)) = 0$.

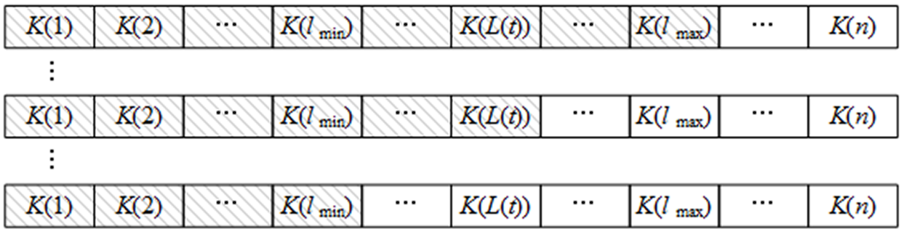


Figure 2. Embedding paths of T stego images embedded secret messages of different lengths, where the squares with oblique lines denote the stego positions.

Therefore, in theory, if one can obtain T stego images embedded along the same path, then the stego positions can be located according to the embedding ratio in each position, $f(i)$; if one can obtain enough stego images embedded messages with all possible lengths along the same path, then the embedding path can be recovered according to the embedding ratio in each position, $f(i)$.

2.2. *Estimation of embedding ratio in each position based on pixel block set*

In practice, steganalysts do not know the embedding ratio in each position, so they must try to estimate it. The idea is to collect all pixels in the same position i of all stego images to form the pixel subset of position i, and then use the quantitative steganalysis algorithm to estimate the embedding ratio of the formed set. However, the existing quantitative steganalysis algorithms usually consider the correlation between multiple pixels. The pixels in the formed subset come from different images, so usually they are not correlated. This means that most of the existing quantitative steganalysis algorithms cannot be directly used to estimate the embedding ratio in each position. Therefore, this subsection presents a method to estimate the embedding ratio in each position based on two categories of pixel block sets.

The two categories of pixel block sets are constructed as shown in Figure 3. For each position i, N adjacent pixels containing the pixel in position i are selected from the t-th stego image to form a pixel block

$$\mathbf{D}_{t,i} = \{s_{t,i}, s_{t,i_1}, s_{t,i_2}, \ldots, s_{t,i_{N-1}}\}. \tag{2.1}$$

Then the pixel $s_{t,i}$ is eliminated from $\mathbf{D}_{t,i}$ to form another block

$$\mathbf{d}_{t,i} = \{s_{t,i_1}, s_{t,i_2}, \ldots, s_{t,i_{N-1}}\}. \tag{2.2}$$

Then, the pixel blocks $\mathbf{D}_{t,i}$ and $\mathbf{d}_{t,i}$ are united over $1 \leq t \leq T$ respectively as follows to obtain the pixel block set containing pixels in position i and the pixel block set not containing pixels in position i:

$$\begin{cases} \mathbf{D}_i = \bigcup_{t=1}^{T} \mathbf{D}_{t,i} \\ \mathbf{d}_i = \bigcup_{t=1}^{T} \mathbf{d}_{t,i} \end{cases} \tag{2.3}$$

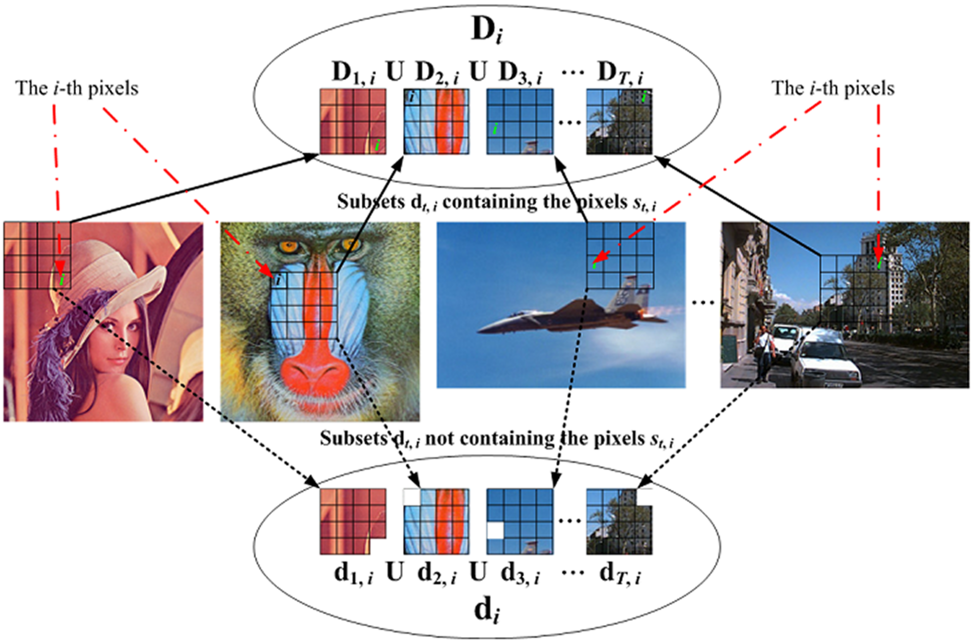

$\mathbf{D}_i$
$\mathbf{D}_{1,i}$ U $\mathbf{D}_{2,i}$ U $\mathbf{D}_{3,i}$ ··· $\mathbf{D}_{T,i}$
The i-th pixels
The i-th pixels
Subsets $\mathbf{d}_{t,i}$ containing the pixels $s_{t,i}$
Subsets $\mathbf{d}_{t,i}$ not containing the pixels $s_{t,i}$
$\mathbf{d}_{1,i}$ U $\mathbf{d}_{2,i}$ U $\mathbf{d}_{3,i}$ ··· $\mathbf{d}_{T,i}$
$\mathbf{d}_i$

Figure 3. Constructing the pixel block set containing the i-th pixels of all stego images and another pixel block set not containing them.

Let P_i denote the ratio of the number of stego pixels (viz. the pixels containing messages) to the number of pixels in $\mathbf{D}_i$, p_i denote the ratio of the number of stego pixels to the number of pixels in $\mathbf{d}_i$. Then, from the definitions of $\mathbf{D}_{t,i}$, $\mathbf{d}_{t,i}$, $\mathbf{D}_i$ and $\mathbf{d}_i$ in formulas (2.1)-(2.3), it can be seen that the difference between the number of stego pixels in $\mathbf{D}_i$ and the number of stego pixels in $\mathbf{d}_i$ is equivalent to the number of stego pixels in the position i of all T stego images.

Therefore, the ratio of the number of stego pixels in the i-th pixels of all T stego images can be obtained as follows,

$$f(i) = \frac{(|\mathbf{D}_i|P_i - |\mathbf{d}_i|p_i)}{T} = NP_i - (N-1)p_i. \tag{2.4}$$

where $|\mathbf{D}_i|$ and $|\mathbf{d}_i|$ denotes the numbers of pixels in $\mathbf{D}_i$ and $\mathbf{d}_i$, respectively.

From (2.4), it can be seen that estimating the embedding ratio in position i can be converted to estimating the embedding ratios in the pixel block sets $\mathbf{D}_i$ and $\mathbf{d}_i$. Because there are some adjacent pixels for each pixel in the pixel block sets $\mathbf{D}_i$ and $\mathbf{d}_i$, if one owns enough stego images embedded along the same path, the existing quantitative steganalysis methods that consider the correlation between adjacent pixels should be able to estimate the embedding ratios in $\mathbf{D}_i$ and $\mathbf{d}_i$, which are referred to as $\hat{P}_i$ and $\hat{p}_i$. Then, the embedding ratio in position i can be estimated as follows,

$$\hat{f}(i) = N\hat{P}_i - (N-1)\hat{p}_i. \tag{2.5}$$

2.3. *Methodology of locating steganalysis*

Because there is a significant difference between the embedding ratios in stego position and non-stego position, and the embedding ratio in position i can be estimated by the method proposed in section 2.2, the following methodology of locating steganalysis can be derived to judge whether the pixels in a position i contain secret messages.

1) Select N adjacent pixels containing the pixel in position i from each stego image to form a pixel block $\mathbf{D}_{t,i}$, then eliminate the pixel in position i to form another pixel block $\mathbf{d}_{t,i}$.

2) Collect the pixel block $\mathbf{D}_{t,i}$ of the position i from T stego images to form the pixel block set $\mathbf{D}_i$, and collect the pixel block $\mathbf{d}_{t,i}$ of the position i from T stego images to form the pixel block set $\mathbf{d}_i$.

3) According to the steganography algorithm used to embed messages in T stego images, adopt a proper quantitative steganalysis algorithm to estimate the embedding ratios in $\mathbf{D}_i$ and $\mathbf{d}_i$, viz. $\hat{P}_i$ and $\hat{p}_i$.

4) Apply the estimated embedding ratios $\hat{P}_i$ and $\hat{p}_i$ to Eqn. (2.5), then obtain the estimated embedding ratio in position i, $\hat{f}(i)$.

5) According to the estimated embedding ratios of all positions, the positions are arranged in descending order. If the estimated embedding ratio $\hat{f}(i)$ is one of the first l positions in the arranged sequence, the position i is regarded as a stego position; otherwise, the position i is regarded as a non-stego position.

3. Locating steganalysis of LSB replacement from structural and WS steganalysis

Although there have been numerous steganography techniques presented, the simple LSB replacement steganography is still a popular one because of its excellent concealment from the human sense system and its extreme simplicity. Many tools in the Internet are still developed based on it. In this paper, two typical structural steganalysis methods and the WS method will be applied to the above methodology to locate the messages embedded using LSB replacement. Whether a given quantitative steganalysis method can be applied to the above methodology to locate the messages is determined by whether it can estimate the embedding ratios in $\mathbf{D}_i$ and $\mathbf{d}_i$ in Figure 3. Therefore, in the following, the applicability of the SPA (Sample Pair Analysis) method, the RS (Regular group and Singular group) method, and the WS method to $\mathbf{D}_i$ and $\mathbf{d}_i$ will be discussed.

3.1. Locating steganalysis of LSB replacement from structural steganalysis

For LSB replacement, structural steganalysis is a main category of quantitative steganalysis methods which have attracted much attention in the past twenty years. The SPA [10] and RS [11] methods are two typical structural steganalysis methods designed for estimating the embedding ratio of LSB replacement.

In the SPA method, the pixel pairs are used as the basic unit to utilize the correlation between adjacent pixels. The SPA method divides the pixel pairs in the given object into many trace sets depending on the absolute difference between two pixel values in a pixel pair. Each trace set is again divided into four different trace subsets depending on the LSBs of the pixel values in a pixel pair. When embedding a message into an image by LSB replacement, each pixel pair should transfer from one of the four trace subsets of the trace set containing it to another. Consequently, the statistical relations between the cardinalities of four trace subsets should be changed. Dumitrescu et al. modelled the relations between the cardinalities of these trace subsets, and derived a quadratic equation to estimate the embedding ratio based on the following assumptions:

(a) The message is randomly embedded into the LSB plane of the cover image, so the LSB of each pixel is altered with the same probability.

(b) For natural images, it is equally probable that the LSB of the larger component of a pixel pair differing by an odd value is 0 or 1.

In the RS method, the image is divided into disjoint groups of adjacent pixels. In typical images, flipping the LSBs of some pixels will lead to an increase in the noisiness rather than a decrease. So, all groups are classified into the following three types depending on whether the noisiness of each group will increase, decrease, or be unchanged after adding a noise. Fridrich et al. defined a multituples as a mask to describe the pattern of noise added to a group. The mask $\mathbf{M}$ composing of 0 and 1 denotes that the noise is added by flipping the LSBs of pixels corresponding to the element 1. The mask $-\mathbf{M}$ where the element 1 in $\mathbf{M}$ is replaced with -1 denotes that the noise is added by adding 1 to the odd pixels and subtracting 1 from the even pixels corresponding to the element -1. Then, Fridrich et al. pointed out that when adding the noise with a mask $\mathbf{M}$ to an image, the relationship between the embedding ratio in the image and the ratio of regular groups (or singular groups) can be well modelled with a straight line; and when adding the noise with a mask $-\mathbf{M}$ to an image, the relationship between the embedding ratio in the image and the ratio of regular groups (or singular groups) can be approximated using a second-degree polynomial. Finally, a quadratic equation was derived to estimate the embedding ratio based on the relationships above and the following assumptions:

(c) For natural images, the ratio of regular groups with mask $\mathbf{M}$ is approximately equal to the ratio of regular groups with mask $-\mathbf{M}$, and the ratio of singular groups with mask $\mathbf{M}$ is approximately equal to the ratio of singular groups with mask $-\mathbf{M}$.

(d) For stego images embedded fully, the ratio of regular groups with positive mask $\mathbf{M}$ is approximately equal to the ratio of singular groups with the same positive mask.

Therefore, if one wants to use the RS and SPA methods to estimate the embedding ratios of LSB replacement in $\mathbf{D}_i$ and $\mathbf{d}_i$ to locate the embedded messages, the selected pixel block sets $\mathbf{D}_i$ and $\mathbf{d}_i$ should satisfy the assumptions (a)~(d).

For the assumption (a), one can randomly select a rectangular pixel block as $\mathbf{D}_{t,i}$ from all possible rectangular pixel blocks with N pixels containing the i-th pixel, then eliminate the i-th pixel to obtain the pixel block $\mathbf{d}_{t,i}$. This would avoid the possible phenomenon that the message bits are not scattered randomly in the selected pixel blocks.

For the assumptions (b) and (c), some experiments were carried out in the following experimental setup to test that whether the selected pixel block sets $\mathbf{D}_i$ and $\mathbf{d}_i$ satisfy them:

3000 images were downloaded from http://photogallery.nrcs.usda.gov, originally very high resolution color images in the format “tiff”. Then for each image, three sub-images of size 512×512 pixels were cut from the top left corner, middle, and bottom right corner. The 9000 images obtained were converted to grayscale images in the format “bmp”. (The tool used was Advanced Batch Converter 3.8.20.) The horizontally and vertically adjacent pixel pairs were utilized. This is a representative option in existing literature.

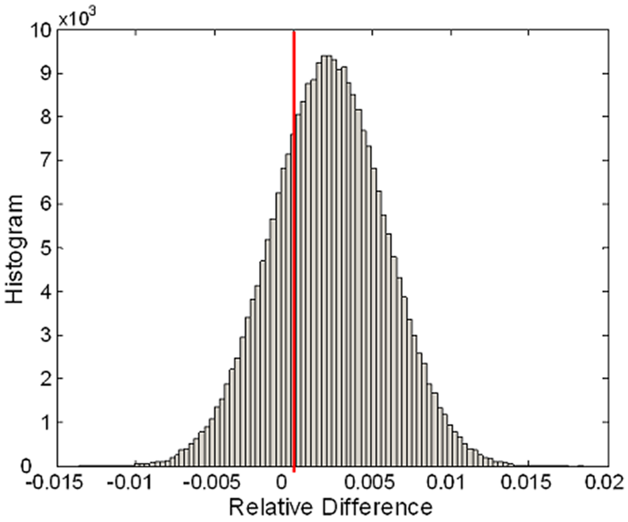


Figure 4. Histogram of the relative differences between Y and X.

Let Y denote the number of pixel pairs whose two pixels differ by an odd value and larger pixel's LSB is 1, X denote the number of pixel pairs whose two pixels differ by an odd value and larger pixel's LSB is 0. Figure 4 shows the histogram of relative differences between Y and X to the total number of pixel pairs differing by an odd value for $\mathbf{D}_i$. For the block set $\mathbf{d}_i$, the histogram is very similar to Figure 4, so it is not displayed. As seen from Figure 4, the absolute values of the relative differences are less than 1% for most block sets $\mathbf{D}_i$, larger than 1% for only a few block sets $\mathbf{D}_i$, and all less than 2%. This demonstrates the rationality of the assumption (b) for $\mathbf{D}_i$ and $\mathbf{d}_i$. Because the equivalence between the assumptions (b) and (c) has been proven in [34], the experimental evidence of assumption (c) is not supplied again.

From the definition of mask, applying positive mask to a group is actually flipping the LSBs of some pixels in the group. In a group of pixels from the fully embedded stego image, the LSBs of the pixels have been randomized. Then, flipping the LSBs of any pixel in this group will increase and decrease the noisiness with the same probability. Further, in the subsets $\mathbf{D}_i$ and $\mathbf{d}_i$, the message bits embedded into the LSBs of the pixels are considered to be pseudo-random. Therefore, the assumption (d) will still hold for the subsets $\mathbf{D}_i$ and $\mathbf{d}_i$.

3.2. Locating steganalysis of LSB replacement from WS steganalysis

The WS method is another typical quantitative steganalysis method for LSB replacement. The WS method performs better than structural steganalysis for large embedding ratios. This can effectively offset the defect of structural steganalysis.

The WS method was first proposed in [16]. Then, Ker [25] described the core point of the method in a slightly different way, making the residuals explicit. For each sample, Ker defined the residual $r_{t,i} = (s_{t,i} - \bar{s}_{t,i})(s_{t,i} - \hat{c}_{t,i})$ where $\bar{s}_{t,i}$ denotes the value of $s_{t,i}$ with LSB flipped, and $\hat{c}_{t,i}$ denotes the estimated cover pixel value responding to $s_{t,i}$. Then, it was derived that the mean residual of all pixels is an unbiased estimation of the modification ratio caused by LSB replacement, viz. the double mean residual is an unbiased estimation of the embedding ratio.

When applying the WS method to estimating the embedding ratio in the subset $\mathbf{D}_i$ and $\mathbf{d}_i$, the estimated embedding ratios are $\hat{P}_i = \frac{2\sum_{s_{t,k}\in\mathbf{D}_i} r_{t,k}}{NT}$ and $\hat{p}_i = \frac{2\sum_{s_{t,k}\in\mathbf{d}_i} r_{t,k}}{(N-1)T}$. Then, from formula (2.5), the estimated embedding ratio in the i-th pixels of all T stego images is

$$\hat{f}(i) = \frac{2\left(\sum\limits_{s_{t,\,k}\in\mathbf{D}_i} r_{t,\,k} - \sum\limits_{s_{t,\,k}\in\mathbf{d}_i} r_{t,\,k}\right)}{T} = \frac{2\sum\limits_{t=1}^{T} r_{t,i}}{T}. \tag{3.1}$$

It can be seen that when applying the WS method to the presented methodology to locate the messages embedded by LSB replacement, the obtained detector is equivalent to the detector in [25]. Therefore, the detector in [25] can be viewed as a special case of the presented methodology.

4. Experimental results and analysis

The locators of LSB replacement from SPA, RS, and WS methods were evaluated on the 9000 images introduced in Section 3.1 when verifying assumption (b) used by the SPA method. Before embedding, a pseudo-random embedding path was generated by permuting the numbers 1, 2, ..., 512×512. At first, for each image in the cover image set, 40960 pseudo-random bits (5K bytes) were embedded into the pixels' LSBs along the pseudo-random embedding path. Then 9000 stego images with 40960 pseudo-random bits in the same positions were obtained to compose the first group of test images. Secondly, the 9000 cover images were partitioned into 4500 cells, and each cell contained 2 images. Then, 8 pseudo-random bits (1 byte), 16 pseudo-random bits (2 bytes), ..., and 36000 pseudo-random bits (4500 bytes) were embedded into the images in the 4500 cells respectively. The 9000 stego images embedded with messages of different lengths composed of the second group of test images. The horizontally and vertically adjacent pixel pairs were utilized in the locator from SPA. In order to be consistent with the locator from SPA, each group contained two adjacent pixels in the locator from RS. In the locator from WS, the arithmetic average of the four closest neighbors of each pixel was taken as the estimator of the cover pixels which was the same as it used in the original WS method [16].

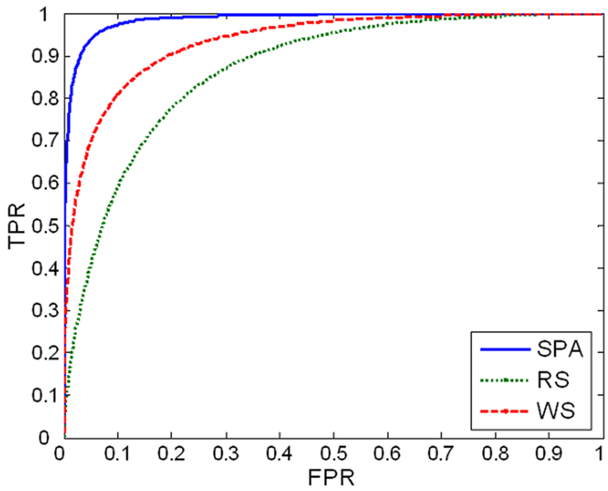


Figure 5. ROC curves for 1000 stego images embedded 40960 bits in the same positions.

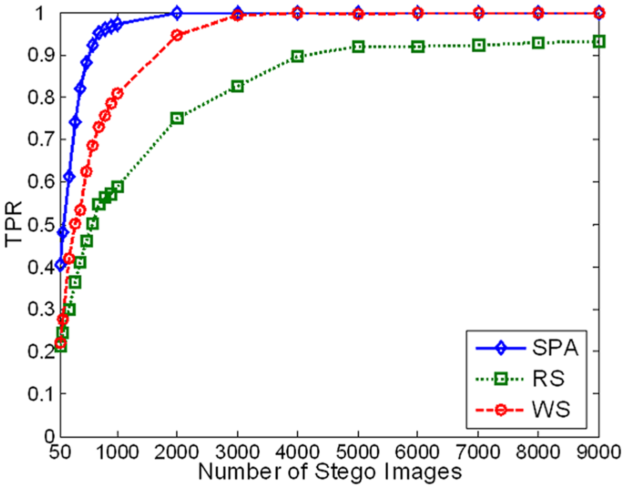
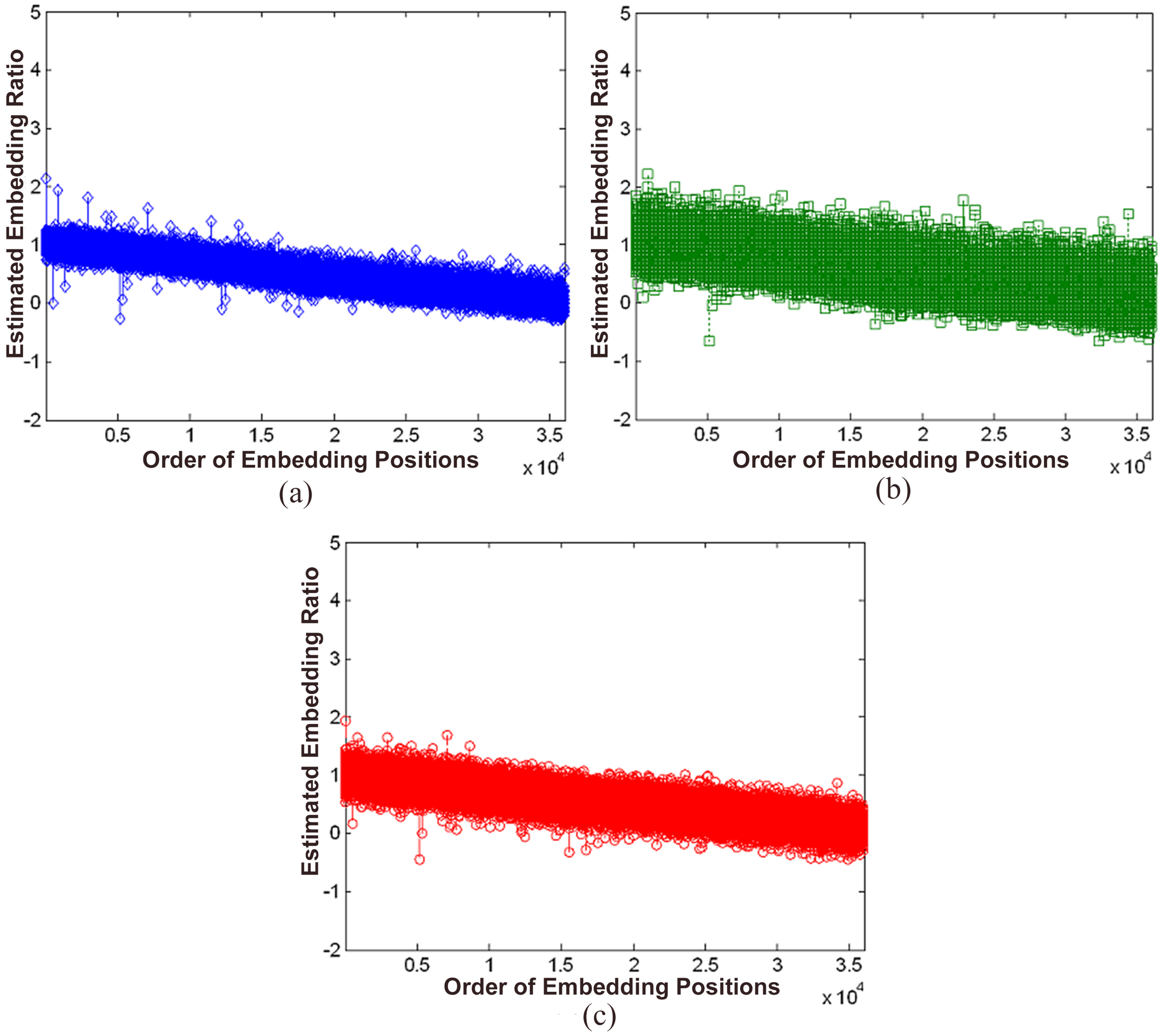


Figure 6. Observed true positive rates when the false positive rate is 10% for the cases of owning 500, 100, 200, . . . , 1000, 2000, . . . , and 9000 stego images.

For evaluating the performances of the locators from SPA, RS, and WS, 1000 stego images were selected from the first group. Figure 5 shows the receiver operating characteristic (ROC) curves for 1000 stego images with 40960 bits in the same positions, which display how the false positive rate (FPR) and true positive rate (TPR) vary as a threshold for the diagnosis of stego positions is altered. Observe that, given 1000 stego images embedded messages in the same positions, the stego positions can be discriminated reliably by the locators from SPA, RS, and WS. Specifically, the locator from SPA can locate the stego positions with the highest reliability.

Ideally, the performance should be tested for different numbers of stego images, but it is impossible to display the ROC curves for every possible number of stego images owned. Therefore, the true positive rates are given for 50, 100, 200, . . . , 1000, 2000, . . . , 9000 stego images when the false positive is 10% (see Figure 6). As can be seen from Figure 6, with an increase in the number of stego images owned, the true positive rates of the three locators increase significantly. Specifically, the locator from SPA can locate more than 90% of the stego positions when the number of stego images owned is more than 500. The locator from WS is second best.

Additionally, because the messages should be embedded into the pixels along the generated pseudo-random embedding path, for the second group of stego images embedded messages of lengths from 8 bits to 3600 bits, the values of the estimated embedding ratios in the first 3600 embedding positions should descend along the generated pseudo-random embedding path. From Figure 7, one can see that the values of the estimated embedding ratios descend in the rough, which is consistent with the analysis. However, because of the errors of quantitative steganalysis, it is still very difficult to estimate the sequence of embedding positions based on the estimated embedding ratio in each position. This further demonstrates that it is very important to design quantitative steganalysis with higher precision.



(a)

(b)

(c)

Figure 7. Estimated embedding ratios in the first 3600 embedding positions for stego images embedded messages of lengths from 8 bits to 3600 bits. (a) Embedding ratios estimated by the locator from SPA. (b) Embedding ratios estimated by the locator from RS. (c) Embedding ratios estimated by the locator from WS.

5. Conclusions

Locating steganalysis can provide important clues for extracting secret messages embedded by stegaography. However, there is still no universal locating steganalysis methodology. By considering the idea of "from blind to quantitative steganalysis", under the condition of owning multiple stego images embedded along the same embedding path, this paper proposes a locating steganalysis methodology by constructing two categories of pixel block sets and using quantitative steganalysis. Then, three locating steganalysis algorithms are designed for LSB replacement steganography by applying three typical structural steganalysis methods: SPA and RS, and weighted stego image steganalysis method WS to the proposed methodology. The experimental results verify the validities of the presented methodology.

The presented methodology points out a new use of quantitative steganalysis, and further demonstrates that it is necessary to design more precise quantitative steganalysis methods.

In the future, we will try to apply more quantitative steganalysis methods to the presented methodology to locate messages embedded by other steganography methods, and fuse multiple quantitative steganalysis methods using machine learning [35] to improve the locating accuracy. Additionally, we may search for images with similar content on the Internet [36], and use them as references to locate the payload of the new steganography algorithm [37].

Acknowledgments

This work is supported by the National Natural Science Foundation of China (Grant Nos. 61872448, U1536104, 61772549, U1736214, 61602508 and 61601517) and Science and Technology Research Project of Henan Province, China (Grant No. 152102210005).

Conflict of interest

The author declares there is no conflict of interest.

References

1. J. Fridrich and M. Golijan, Practical steganalysis of digital images—state of the art, *Proc. SPIE*,**4675** (2002), 1–13.
2. L. Xiang, Y. Li, W. Hao, et al., Reversible natural language watermarking using synonym substitution and arithmetic coding, *Comput. Mater. Con.*, **55** (2018), 541–559.
3. Y. Zhang, C. Qin, W. Zhang, et al., On the fault-tolerant performance for a class of robust image steganography, *Signal Process.*, **146** (2018), 99–111.
4. Y. Zhang, D. Ye, J. Gan, et al., An image steganography algorithm based on quantization index modulation resisting scaling attacks and statistical detection, *Comput. Mater. Con.*, **56** (2018), 151–167.
5. J. Chen, W. Lu, Y. Yeung, et al., Binary image steganalysis based on distortion level co-occurrence matrix, *Comput. Mater. Con.*, **55** (2018), 201–211.
6. Y. Ma, X. Luo, X. Li, et al., Selection of rich model steganalysis features based on decision rough set α-positive region reduction, *IEEE Trans. Circ. Syst Vid. Technol.*, **29** (2019), 336–350.
7. J. Fridrich, M. Goljan and D. Soukal, Searching for the stego-key, *Proc. SPIE*, (2004), 70–82.
8. J. Fridrich, M. Goljan, D. Soukal, et al., Forensic steganalysis: determining the stego key in spatial domain steganography, *Proc. SPIE*, (2005), 631–642.
9. J. Fridrich, M. Goljan, D. Hogea, et al., Quantitative steganalysis of digital images: estimating the secret message length, *ACM Mult. Syst. J.*, **9** (2003), 288–302.
10. S. Dumitrescu, X. Wu and Z. Wang, Detection of LSB steganography via sample pair analysis, *IEEE Trans. Signal Proces.*, **51** (2003), 1995–2007.

11. J. Fridrich, M. Goljan and R. Du, Detecting LSB steganography in color and gray-scale images, *IEEE Mult.*, **8** (2001), 22–28.

12. A. D. Ker, A general framework for the structural steganalysis of LSB replacement, in *Proc. Information Hiding* (eds. M. Barni, J. Herrera-Joancomartí, S. Katzenbeisser, F. Pérez-González), Springer-Verlag, (2005), 296–311.

13. C. Yang, F. Liu, X. Luo, et al., Steganalysis frameworks of embedding in multiple least-significant bits, *IEEE. Trans. Inf. Foren. Sec.*, **3** (2008), 662–672.

14. J. Kodovský and J. Fridrich, Quantitative structural steganalysis of JSteg, *IEEE. Trans. Inf. Foren. Sec.*, **5** (2010), 681–693.

15. C. Yang, F. Liu, X. Luo, et al., Pixel group trace model-based quantitative steganalysis for multiple least-significant bits steganography, *IEEE. Trans. Inf. Foren. Sec.*, **8** (2013), 216–228.

16. J. Fridrich and M. Goljan, On estimation of secret message length in LSB steganography in spatial domain, *Proc. SPIE*, (2004), 23–34.

17. T. Pevný, J. Fridrich and A. D. Ker, From blind to quantitative steganalysis, *Proc. SPIE*, (2009), 72540C.

18. Q. Guan, J. Dong and T. Tan, Blind quantitative steganalysis based on feature fusion and gradient boosting, in *Proc. Int. Workshop Digit. Watermark* (eds. H. J. Kim, Y. Q. Shi and M. Barni), Springer-Verlag, (2011), 266–279.

19. T. T. Quach, Locating payload embedded by group-parity steganography, *Digit. Invest.*, **9** (2012), 160–166.

20. T. T. Quach, Extracting hidden messages in steganographic images, *Digit. Invest.*, **11** (2014), S40–S45.

21. J. Liu, Y. Tian, T. Han, et al., Stego key searching for LSB steganography on JPEG decompressed image, *Sci. China Inform. Sci.*, **59** (2016), 32105.

22. A. Westfeld and A. Pfitzmann, Attacks on steganographic systems, in *Proc. Inform. Hid.* (eds. A. Pfitzmann), Springer-Verlag, (2000), 61–75.

23. S. Trivedi and R. Chandramouli, Secret key estimation in sequential steganography, *IEEE. Trans. Singal Proces.*, **53** (2005), 746–757.

24. A. D. Ker and R. Böhme, Revisiting weighted stego-image steganalysis, in *Proc. SPIE* (eds. E. J. Delp, P. W. Wong, J. Dittmann and N. D. Memon), SPIE, (2008), 681905.

25. A. D. Ker, Locating steganographic payload via WS residual, in *Proc. ACM Mult. Sec.* (eds. A. D. Ker, J. Dittmann and J. Fridrich), ACM, (2008), 27–32.

26. A. D. Ker and I. Lubenko, Feature reduction and payload location with wam steganalysis, in *Proc. SPIE* (eds. E. J. Delp, J. Dittmann, N. D. Memon and P. W. Wong), SPIE, (2009), 72540A.

27. Y. Luo, X. Li and B. Yang, Locating steganographic payload for LSB matching embedding, in *Proc. ICME* (eds. I. Cheng, G. Fernandez and H. Wang), IEEE, (2011), 1–6.

28. T. T. Quach, On locating steganographic payload using residuals, in *Proc. SPIE* (eds. N. D. Memon, J. Dittmann, A. M. Alattar and E. J. Delp), SPIE, 2011, 0J1–0J7.

29. T. T. Quach, Optimal cover estimation methods and steganographic payload location, *IEEE Trans. Inf. Foren. Sec.*, **6** (2011), 1214–1222.

30. X. Gui, X. Li and B. Yang, Improved payload location for LSB matching steganography, in *IEEE ICIP* (eds. E. Saber, S. Hemami and G. Sharma), IEEE, (2012), 1125–1128.

31. T. T. Quach, Cover estimation and payload location using Markov random fields, in *Proc. SPIE* (eds. A. M. Alattar, N. D. Memon and C. D. Heitzenrater), SPIE, (2014), 90280H.

32. J. Liu, Y. Tian, T. Han, et al., LSB steganographic payload location for JPEG-decompressed images, *Digit. Signal Process*, **38** (2015), 66–76.

33. C. Yang, X. Luo, J. Lu, et al., Extracting hidden messages of MLSB steganography based on optimal stego subset, *Sci. China Inform. Sci.*, **61** (2018), 119103:1–119103:3.

34. X. Luo, C. Yang and F. Liu, Equivalence analysis among DIH, SPA, and RS steganalysis methods, in *Proc. IFIP CMS* (eds. H. Leitold and E. P. Markatos), Springer-Verlag, (2006), 161–172.

35. L. Xiang, G. Zhao, Q. Li, et al., TUMK-ELM: a fast unsupervised heterogeneous data learning approach, *IEEE Access*, **6** (2018), 35305–35315.

36. L. Xiang, X. Shen, J. Qin, et al.,Discrete multi-graph hashing for large-scale visual search, *Neural Process Lett.*, (2018).

37. L. Liu, Z. Wang, Z. Qian, et al.,Steganography in beautified images, *Math. Biosci. Engineer.*, **16** (2019), 2333–2333.

AIMS Press