

Weighted Stego-Image Steganalysis of Messages Hidden into Each Bit Plane

CHUNFANG YANG^{1,*}, FENLIN LIU¹, SHIGUO LIAN², XIANGYANG LUO¹
AND DAOSHUN WANG³

¹Zhengzhou Information Science and Technology Institute, Zhengzhou 450002, China

²Corporate Research, Huawei Technologies, Beijing 100034, China

³Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China

*Corresponding author: chunfangyang@126.com

For hiding messages into multiple least significant bit (MLSB) planes, a new weighted stego-image (WS) steganalysis method is proposed to estimate the ratio of messages hidden into each bit plane. First, a new WS with multiple weights is constructed, and it is proved that when the squared Euclidean distance between the WS and the cover image is minimal, the weight parameters are equal to the embedding ratios in MLSB planes. Afterward, based on this result and an estimation of cover image, a simple estimation equation is derived to estimate the embedding ratio in each bit plane. Experimental results show that the new steganalysis method performs more stably with the change of embedding ratios than typical structural steganalysis, and outperforms the typical structural steganalysis method on the estimation accuracy when the embedding ratio in any bit plane is larger than 0.4.

Keywords: steganography; steganalysis; MLSB; weighted stego-image

Received 30 June 2011; revised 26 August 2011

Handling editor: Wojciech Mazurczyk

1. INTRODUCTION

Steganography is the art of hiding the very presence of communication by embedding secret messages into innocuous looking covers, such as digital images [1]. As an increasing amount of data stored on computers and transmitted over networks, steganography has been one of the key technologies of multimedia information security [2, 3]. Simultaneously, steganography also has a number of nefarious applications, such as hiding records of illegal activity, financial fraud, industrial espionage and communication amongst members of criminal or terrorist organizations [4]. These may cause tremendous socio-economic impact on individuals, enterprises and even countries; thus steganography has been a serious challenge to the digital forensics. Therefore, steganalysis (which is the opposite technology of steganography) has attracted great attentions and been one of the key technologies of digital forensics.

Technically, steganography is considered broken when the mere presence of secret message can be established [1]. However, the digital forensics investigators can use the steganalysis technologies not only to detect the hidden records of illegal activities and the secret communication amongst members of criminal or terrorist organizations, but also to

estimate the length of secret message or the modification ratio of the cover signal [5]. The estimation of the secret message's length or the modification ratio can not only be used to distinguish the stego objects, but also help in the estimation of stego positions and the search of stego key [6–8]. Steganalysis that can estimate the length of the secret message or the modification ratio of the signal samples is called as quantitative steganalysis [5].

Nowadays, amongst numerous steganography methods, replacement of bit planes is a category of important steganography methods because of its simple implementation. The widespread availability of steganography tools developed based on these methods has led to an increased interest in steganalysis techniques for them. For the replacement of LSB (least significant bit) plane, many relevant quantitative steganalysis methods have been designed, such as RS (regular and singular groups) method [9], DIH (difference image histogram) method [10], SPA (sample pair analysis) method [11], WS (weighted stego-image) method [12] and some improved variant [13–15] of these steganalysis methods. But all of above were just designed for LSB replacement specifically; it cannot be expected that they can give correct answers for the size of payload in LSB plane when embedding is also carried

out in the other-LSB planes ([16] has given the reason for two least significant bits (2LSB) embedding). Some relevant works have reported that secret message can not only be embedded into LSB plane, but also be embedded into multiple least significant bit (MLSB) planes in [17, 18].

For the replacement of MLSB planes, there have been some steganalysis methods proposed for two distinct MLSB steganography paradigms—TMSB steganography (‘T’ means this embedding paradigm is typical.) and IMLSB steganography (‘I’ means the messages are embedded into each bit plane independently.), which embed equal ratio of message bits into MLSB planes. And these steganalysis methods for MLSB steganography can mainly be classified as two categories: structural steganalysis [16, 19–21] and WS steganalysis [22–24]. However, the embedders likely embed message into different bit planes with different ratios independently, and this category of MLSB steganography is called as ID-MLSB steganography (where ‘ID’ denotes that the message lengths may be embedded into different bit planes independently with different ratios). Furthermore, some other steganography methods can be classified into the ID-MLSB steganography, such as the adaptive steganography in [18]. Therefore, the quantitative steganalysis of ID-MLSB steganography should be very important to the forensics steganalysis. For this category of MLSB steganography, the steganalysis methods above all fail to estimate the embedding ratio in each bit plane. And so far, only the literature [25] presented a method to estimate the embedding ratio in each bit plane based on the SPA model for MLSB replacement steganography (MSPA model) in [21]. The obtained method was called IDMSPA method, and would be called ID3SPA method when being used to estimate the embedding ratios in the three LSB planes. In [26], the IDMSPA method also has been adapted for the quantitative steganalysis of a category of adaptive steganography that embeds message bits into different bit planes with different ratios based on the block’s noise level of MLSB planes.

Although the methods in [25, 26] can estimate the embedding ratio in each bit plane of MLSB planes, they all own the inherent defect—their performances degrade rapidly with the increase of embedding ratio in any bit plane, and when the embedding ratio in any stego bit plane is middle or large, they will fail to estimate the embedding ratios accurately. Contrarily, the WS steganalysis usually performs more stably and owns good performance for the case of a high embedding ratio [21]. Therefore, we try to estimate the embedding ratios of ID-MLSB steganography based on WS more stably and accurately than the existing methods. The main results of this paper are as follows.

- (i) A new WS with multiple weight parameters is defined for ID-MLSB steganography.
- (ii) A theorem is proved that when the squared Euclidean distance between the WS and the cover image is minimal, the obtained weight parameters are equal to the embedding ratios in the MLSB planes.

- (iii) Based on this theorem and an estimated cover image, a quantitative steganalysis method is proposed for estimating the ratio of message hidden into each bit plane. Experimental results show that the new steganalysis method performs more stably with the change of embedding ratio than typical structural steganalysis, and outperforms the structural steganalysis on estimation accuracy when the embedding ratio in any bit plane is middling or large.

The remainder of this paper is organized as follows. Section 2 introduces the related works, including the WS steganalysis for LSB steganography and the ID-MLSB steganography briefly. Section 3 describes the proposed weighted stego steganalysis for ID-MLSB steganography. Finally, a series of experimental results are given in Section 4. The paper closes in Section 5 with the conclusions.

2. RELATED WORKS

2.1. WS steganalysis for LSB steganography

The WS steganalysis was proposed by Fridrich and Goljan [12], and occupies an unusual position in the steganalysis of LSB replacement. Unlike the structural detectors, this method does not use the pixel group analysis on which almost every other reasonably accurate detector relies, but has fairly good accuracy; moreover, it retains its estimation accuracy when embedding changes are not distributed evenly over the cover [15]. In the following, the basic principle of WS steganalysis for LSB replacement will be introduced briefly.

In [12], let $X = \{x_i\}_{i=1}^n$ be a set of integers in the range $[0, 255]$ representing a grayscale cover image whose size is $n = M \times N$. The value of x_i after flipping its LSB will be denoted as \bar{x}_i , viz. $\bar{x}_i = x_i + 1 - 2(x_i \bmod 2)$. Let $S = \{s_i\}_{i=1}^n$ denote the stego-image after embedding pn ($0 \leq p \leq 1$) bits using LSB replacement in pn pixels randomly selected from the cover image X . Fridrich and Goljan defined the WS with weight parameter q as follows:

$$\begin{cases} S^{(q)} = \{s_i^{(q)}\}, \\ s_i^{(q)} = s_i + \frac{q}{2}(\bar{s}_i - s_i), \end{cases} \quad (1)$$

where $0 \leq q \leq 1$, $i = 1, 2, \dots, n$. Then, Fridrich and Goljan proved that the weight parameter q and the embedding ratio p satisfy the following relationship:

$$p = \arg \min_q \frac{1}{n} \sum_{i=1}^n (s_i^{(q)} - x_i)^2, \quad (2)$$

which shows that $S^{(p)}$ is the closest WS to the cover image in the least square sense among all WSs $S^{(q)}$ for $0 \leq q \leq 1$. Based on this, Fridrich and Goljan formulated the procedure of estimating the unknown embedding ratio p from the stego-image as a

minimization problem as follows:

$$\begin{aligned} \hat{p} &= \arg \min_q \frac{1}{n} \sum_{i=1}^n (s_i^{(q)} - \hat{x}_i)^2 \\ &= -\frac{2}{n} \sum_{i=1}^n (s_i - \hat{x}_i)(1 - 2(s_i \bmod 2)), \end{aligned} \quad (3)$$

where \hat{x}_i denotes the prediction of x_i .

So far the WS steganalysis has been enhanced from many aspects and applied to the steganalysis of some other steganography, such as TMLSB and JSteg steganography.

2.2. ID-MLSB steganography

When one wants to replace the bits in the l (l denotes the number of stego least-significant bit planes) LSB planes with the message bits, at least the following three methods can be adopted:

- (i) TMLSB steganography, which randomly selects fixed number of pixels and replaces their l LSBs with the message fragments with size of l (seeing Fig. 1a for $l = 3$);
- (ii) IMLSB steganography, which randomly selects fixed number of bits from the l LSB planes and replaces them with the message bits (seeing Fig. 1b for $l = 3$);
- (iii) ID-MLSB steganography, for each bit plane of the MLSB planes, which randomly selects fixed number of bits from the bit plane, and replaces them with the message bits (seeing Fig. 1c for $l = 3$).

From Fig. 1, it can be seen that being different from the TMLSB and IMLSB steganography, the ID-MLSB steganography possibly embeds message bits into different bit planes with different ratios. Furthermore, some other steganography methods can be classified into the ID-MLSB steganography, such as the adaptive steganography in [18]. The main steps of ID-MLSB steganography can be described as follows:

- (i) Determine the number of LSB planes that will contain the message bits, viz. l ;
- (ii) For each bit plane in the l LSB planes, randomly select fixed number of bit positions from the bit plane;
- (iii) If the bit in the selected position is different from the message bit, flip the bit in the selected position; otherwise, do not change the bit.

3. PROPOSED WS STEGANALYSIS FOR ID-MLSB STEGANOGRAPHY

3.1. WS for ID-MLSB steganography

In ID-MLSB steganography, the messages hidden into l LSB planes may own different lengths independently. Thus, let $\mathbf{p} = (p_1, p_2, \dots, p_l)$ denote the vector of embedding ratios where

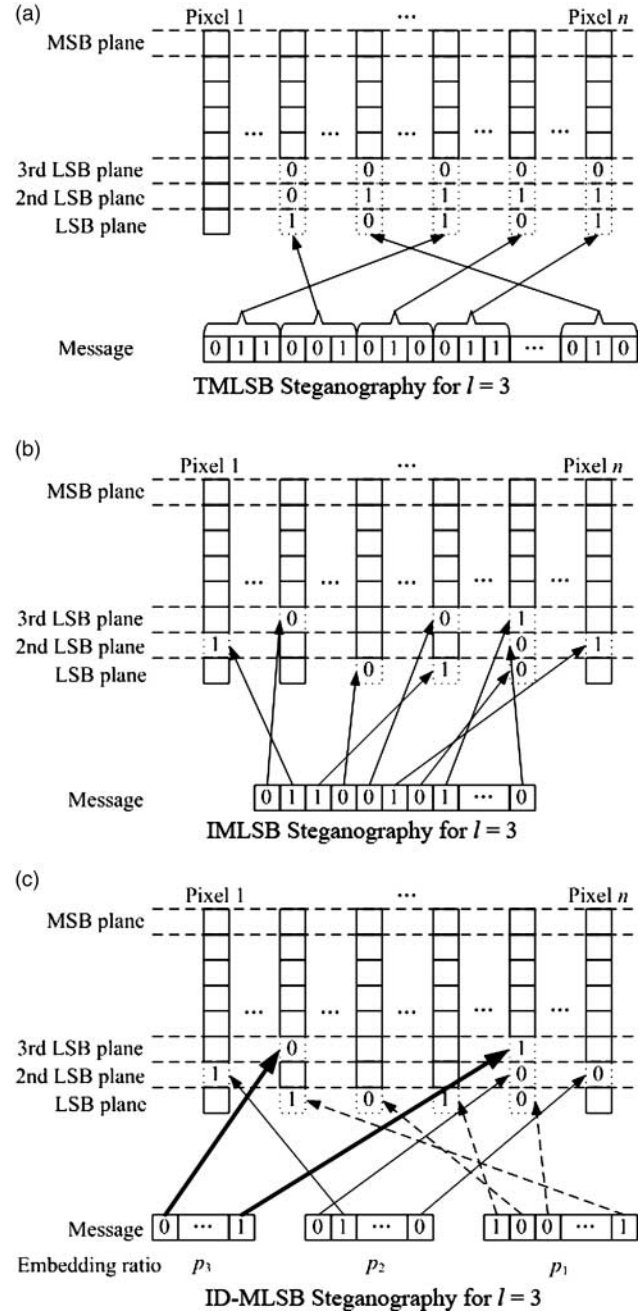


FIGURE 1. Comparison among ID-MLSB steganography and other two categories of MLSB steganography. (a) TMLSB steganography for $l = 3$. (b) IMLSB steganography for $l = 3$. (c) ID-MLSB steganography for $l = 3$.

p_i ($1 \leq i \leq l$) is the embedding ratio in the i th LSB plane. Then the WS for ID-MLSB steganography can be defined as follows:

$$\begin{cases} S^{(q,l)} = \{s_i^{(q,l)}\}, \\ s_i^{(q,l)} = s_i + \sum_{j=1}^l \frac{q_j}{2} [2^{j-1} - 2(s_i \oplus 2^{j-1})], \end{cases} \quad (4)$$

where $\mathbf{q} = (q_1, q_2, \dots, q_l)$ is the vector of weight parameters, and \oplus is the operation of exclusive OR between the operands. The WS in (4) actually reflects the following means:

- (i) If the j th LSB of a pixel is 1, this bit will be subtracted a weight $q_j/2$ from;
- (ii) If the j th LSB of a pixel is 0, this bit will be added a weight $q_j/2$ to.

For the j th LSB of a pixel in the stego-image with embedding ratio $\mathbf{p} = (p_1, p_2, \dots, p_l)$, four cases may occur in it. If the j th LSB of the pixel in the stego image is 1, one of the following two cases must occur:

- (i) The j th LSB of the pixel in the cover image is 1, and the embedding did not change it;
- (ii) The j th LSB of the pixel in the cover image is 0, and the embedding added 1 to it.

If the j th LSB of the pixel in the stego image is 0, one of the other two cases must occur:

- (iii) The j th LSB of the pixel in the cover image is 0, and the embedding did not change it;
- (iv) The j th LSB of the pixel in the cover image is 1, and the embedding subtracted 1 from it.

Because it is approximately equally probable that the l LSBs of a pixel is 0, 1, \dots , $2^l - 1$, the following probabilities can be obtained:

- (i) The probability that the j th LSB is obtained from unchanging is $(1 - p_j/2)$;

- (ii) The probability that the odd j th LSB is obtained from adding 1 to the cover bit is $p_j/2$;
- (iii) The probability that the even j th LSB is obtained from subtracting 1 from the cover bit is $p_j/2$.

Therefore, the construction of WS can be regarded as the procedure of estimating the cover image on the average by inverted operations, viz. when the j th LSB of a pixel is even, adding a weight $q_j/2$ to the j th LSB, otherwise, subtracting a weight $q_j/2$ from the j th LSB.

3.2. Estimating embedding ratio for each bit plane

Because the WS can be regarded as an estimation of the cover image on the average, in theory, when the WS is closest to the cover image, the corresponding weight parameters should be equal to the embedding ratios. Based on this idea, the following theorem will formalize the quantitative steganalysis of ID-MLSB steganography, viz. estimating the embedding ratio in each bit plane of l LSB planes, as a minimization problem (seeing Fig. 2). And then the estimation equation of the embedding ratio in each bit plane will be derived from the theorem and an estimated cover image. Figure 2 shows the principle of estimating embedding ratio in each bit plane based on WS, where MSB is the acronym of most significant bit.

THEOREM 3.1. *If the stego-image $S = \{s_i\}_{i=1}^n$ is obtained by embedding $p_1n, p_2n, \dots, p_l n$ ($0 \leq p_1, p_2, \dots, p_l \leq 1$) random bits into the LSB plane, 2nd LSB plane, \dots , and l th LSB plane of the cover image X respectively using ID-MLSB steganography, then the weight parameters $\mathbf{q} = (q_1, q_2, \dots, q_l)$ and the embedding ratios $\mathbf{p} = (p_1, p_2, \dots, p_l)$*

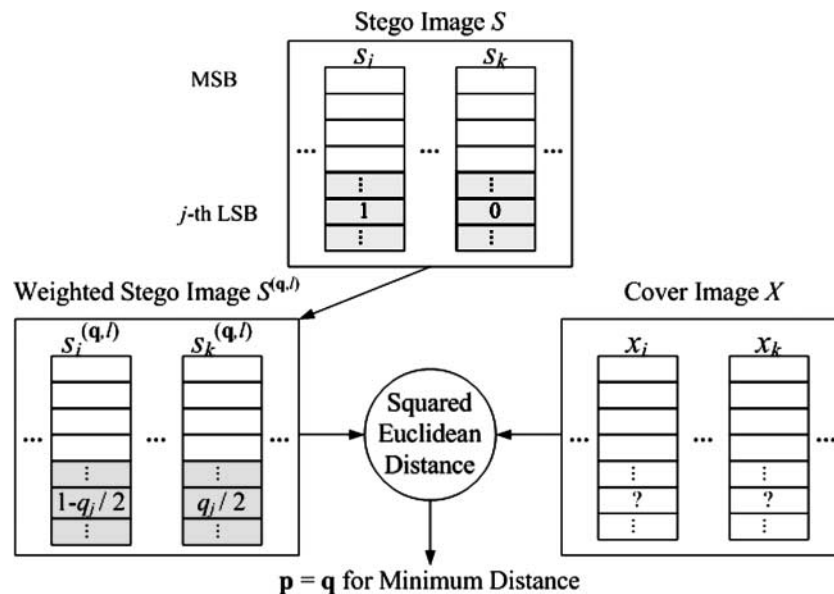


FIGURE 2. Block diagram of the proposed WS steganalysis for ID-MLSB steganography (In practice, the cover image will be replaced by an estimated cover image).

satisfy the following relationship:

$$\begin{cases} \mathbf{p} = \arg \min_{\mathbf{q}} d_l(\mathbf{q}), \\ d_l(\mathbf{q}) = \sum_{i=1}^n (s_i^{(\mathbf{q},l)} - x_i)^2, \end{cases} \quad (5)$$

where $d_l(\mathbf{q})$ is the squared Euclidean distance between the WS and the cover image.

Proof. Applying the definition of WS in (4) to $d_l(\mathbf{q})$, it follows that

$$\begin{aligned} d_l(\mathbf{q}) &= \sum_{i=1}^n (s_i^{(\mathbf{q},l)} - x_i)^2 \\ &= \sum_{i=1}^n \left\{ s_i - x_i + \sum_{j=1}^l \frac{q_j}{2} [2^{j-1} - 2(s_i \oplus 2^{j-1})] \right\}^2 \\ &= \sum_{i=1}^n (s_i - x_i)^2 \\ &\quad + \sum_{i=1}^n \left\{ (s_i - x_i) \sum_{j=1}^l q_j [2^{j-1} - 2(s_i \oplus 2^{j-1})] \right\} \\ &\quad + \sum_{i=1}^n \left\{ \sum_{j=1}^l \frac{q_j}{2} [2^{j-1} - 2(s_i \oplus 2^{j-1})] \right\}^2. \end{aligned} \quad (6)$$

Let

$$\begin{cases} A_l = \sum_{i=1}^n (s_i - x_i)^2, \\ B_l = \sum_{i=1}^n \left\{ (s_i - x_i) \sum_{j=1}^l q_j [2^{j-1} - 2(s_i \oplus 2^{j-1})] \right\}, \\ C_l = \sum_{i=1}^n \left\{ \sum_{j=1}^l \frac{q_j}{2} [2^{j-1} - 2(s_i \oplus 2^{j-1})] \right\}^2, \end{cases} \quad (7)$$

then

$$d_l(\mathbf{q}) = A_l + B_l + C_l. \quad (8)$$

Because the ID-MLSB steganography modifies only the l LSB planes, the $(b-l)$ most significant bits of s_i and x_i are equal. Let $s_{i,j}$ and $x_{i,j}$ denote the j th LSBs of pixels s_i and x_i , respectively. Then,

$$s_i - x_i = \sum_{j=1}^l 2^{j-1} (s_{i,j} - x_{i,j}). \quad (9)$$

Applying $s_{i,j}$ and formula (9) to (7), the formulas of A_l , B_l and C_l become

$$\begin{cases} A_l = \sum_{i=1}^n \left[\sum_{j=1}^l 2^{j-1} (s_{i,j} - x_{i,j}) \right]^2, \\ B_l = \sum_{i=1}^n \left\{ \left[\sum_{k=1}^l 2^{k-1} (s_{i,k} - x_{i,k}) \right] \right. \\ \quad \left. \times \left[\sum_{j=1}^l q_j 2^{j-1} (1 - 2s_{i,j}) \right] \right\}, \\ C_l = \sum_{i=1}^n \left\{ \sum_{j=1}^l q_j 2^{j-2} (1 - 2s_{i,j}) \right\}^2. \end{cases} \quad (10)$$

From the polynomial theorem, the formula of A_l can be written as

$$\begin{aligned} A_l &= \sum_{i=1}^n \left[\sum_{j=1}^l 2^{j-1} (s_{i,j} - x_{i,j}) \right]^2 \\ &= \sum_{i=1}^n \left[\sum_{j=1}^l 2^{2j-2} (s_{i,j} - x_{i,j})^2 \right. \\ &\quad \left. + 2 \sum_{j,k=1,2,\dots,l, j \neq k} 2^{j+k-2} (s_{i,j} - x_{i,j})(s_{i,k} - x_{i,k}) \right] \\ &= \sum_{i=1}^n \sum_{j=1}^l 2^{2j-2} (s_{i,j} - x_{i,j})^2 \\ &\quad + 2 \sum_{i=1}^n \sum_{j,k=1,2,\dots,l, j \neq k} 2^{j+k-2} (s_{i,j} - x_{i,j})(s_{i,k} - x_{i,k}) \\ &= \sum_{j=1}^l \left[2^{2j-2} \sum_{i=1}^n (s_{i,j} - x_{i,j})^2 \right] \\ &\quad + 2 \sum_{j,k=1,2,\dots,l, j \neq k} \left[2^{j+k-2} \sum_{i=1}^n (s_{i,j} - x_{i,j})(s_{i,k} - x_{i,k}) \right]. \end{aligned} \quad (11)$$

Similarly,

$$\begin{aligned} B_l &= \sum_{i=1}^n \left\{ \left[\sum_{k=1}^l 2^{k-1} (s_{i,k} - x_{i,k}) \right] \left[\sum_{j=1}^l q_j 2^{j-1} (1 - 2s_{i,j}) \right] \right\} \\ &= \sum_{j=1}^l \sum_{k=1}^l \sum_{i=1}^n [q_j 2^{j+k-2} (1 - 2s_{i,j})(s_{i,k} - x_{i,k})] \\ &= \sum_{j=1}^l \left[q_j 2^{2j-2} \sum_{i=1}^n (1 - 2s_{i,j})(s_{i,j} - x_{i,j}) \right] \end{aligned}$$

$$+ \sum_{j,k=1,2,\dots,l,j \neq k} \left[q_j 2^{j+k-2} \sum_{i=1}^n (1 - 2s_{i,j})(s_{i,k} - x_{i,k}) \right], \quad (12)$$

$$\begin{aligned} C_l &= \sum_{i=1}^n \left\{ \sum_{j=1}^l q_j 2^{j-2} (1 - 2s_{i,j}) \right\}^2 \\ &= \sum_{i=1}^n \left\{ \sum_{j=1}^l [q_j^2 2^{2j-4} (1 - 2s_{i,j})^2] \right. \\ &\quad \left. + \sum_{j,k=1,2,\dots,l,j \neq k} [q_j q_k 2^{j+k-4} (1 - 2s_{i,j})(1 - 2s_{i,k})] \right\} \\ &= \sum_{j=1}^l \left[q_j^2 2^{2j-4} \sum_{i=1}^n (1 - 2s_{i,j})^2 \right] \\ &\quad + \sum_{j,k=1,2,\dots,l,j \neq k} \left[q_j q_k 2^{j+k-4} \sum_{i=1}^n (1 - 2s_{i,j})(1 - 2s_{i,k}) \right]. \end{aligned} \quad (13)$$

Because it is about equal probable that the l LSBs of the pixel x_i is 0, 1, ..., or $2^l - 1$ [22], it is also about equal probable that the j th ($1 \leq k \leq l$) LSB of the pixel x_i is 0 or 1. And because the embedding ratio in the j th LSB plane of the stego-image is p_j , the probability that the j th LSB is not changed, viz. $s_{i,j} - x_{i,j} = 0$, is $1 - p_j/2$, and the probability that 1 is added to (or 1 is subtracted from) the j th LSB, viz. $s_{i,j} - x_{i,j} = 1$ (or -1), is $p_j/4$. Therefore, the following results can be derived for $j, k = 1, 2, \dots, l$, and $j \neq k$:

$$\sum_{i=1}^n (s_{i,j} - x_{i,j})^2 = \frac{p_j n}{2}, \quad (14)$$

$$\begin{aligned} &\sum_{i=1}^n (s_{i,j} - x_{i,j})(s_{i,k} - x_{i,k}) \\ &= n \times \frac{p_j}{4} \times \frac{p_k}{4} \times [1 \times 1 + (-1) \times 1 \\ &\quad + 1 \times (-1) + (-1) \times (-1)] \\ &= 0, \end{aligned} \quad (15)$$

$$\begin{aligned} &\sum_{i=1}^n (1 - 2s_{i,j})(s_{i,j} - x_{i,j}) \\ &= \frac{p_j n}{4} \times [(-1) \times 1 + 1 \times (-1)] \\ &= -\frac{p_j n}{2}, \end{aligned} \quad (16)$$

$$\begin{aligned} &\sum_{i=1}^n (1 - 2s_{i,j})(s_{i,k} - x_{i,k}) \\ &= \left[\frac{1}{2} \left(1 - \frac{p_j}{2} \right) + \frac{p_j}{4} \right] \frac{p_k n}{4} [1 \times 1 + 1 \times (-1)] \end{aligned}$$

$$+ (-1) \times 1 + (-1) \times (-1)] = 0. \quad (17)$$

Applying (14)–(17) to the formulas of A_l , B_l and C_l in (11)–(13), they can be written as follows:

$$\begin{aligned} A_l &= \sum_{j=1}^l \left(2^{2j-2} \frac{p_j n}{2} \right) \\ &\quad + 2 \sum_{j,k=1,2,\dots,l,j \neq k} (2^{j+k-2} \times n \times 0) \\ &= \frac{n}{2} \sum_{j=1}^l (2^{2j-2} p_j), \end{aligned} \quad (18)$$

$$\begin{aligned} B_l &= \sum_{j=1}^l \left[q_j 2^{2j-2} \times \left(-\frac{p_j n}{2} \right) \right] \\ &\quad + \sum_{j,k \in \{1,2,\dots,l\}, j \neq k} (q_j 2^{j+k-2} \times 0) \\ &= -\frac{n}{2} \sum_{j=1}^l (2^{2j-2} p_j q_j), \end{aligned} \quad (19)$$

$$\begin{aligned} C_l &= \sum_{j=1}^l (q_j^2 2^{2j-4} n) \\ &\quad + \sum_{j,k=1,2,\dots,l,j \neq k} (q_j q_k 2^{j+k-4} \times 0) \\ &= \frac{n}{4} \sum_{j=1}^l (2^{2j-2} q_j^2). \end{aligned} \quad (20)$$

Then, applying (18)–(20) to the Equation (8), the formula of $d_l(\mathbf{q})$ can be written as

$$\begin{aligned} d_l(\mathbf{q}) &= A_l + B_l + C_l \\ &= \frac{n}{2} \sum_{j=1}^l (2^{2j-2} p_j) - \frac{n}{2} \sum_{j=1}^l (2^{2j-2} p_j q_j) \\ &\quad + \frac{n}{4} \sum_{j=1}^l (2^{2j-2} q_j^2). \end{aligned} \quad (21)$$

The partial derivative of $d_l(\mathbf{q})$ with respect to q_k is as follows:

$$\frac{\partial d_l(\mathbf{q})}{\partial q_k} = -2^{2k-3} n p_k + 2^{2k-3} n q_k. \quad (22)$$

From (22), the second partial derivative of $d_l(\mathbf{q})$ with q_k is $2^{2k-3} n$, which is always larger than 0. And because the function $d_l(\mathbf{q})$ is differentiable for $0 \leq q_1, q_2, \dots, q_l \leq 1$, when the partial derivative of $d_l(\mathbf{q})$ with respect to q_1, q_2, \dots, q_l are all zero, the value of $d_l(\mathbf{q})$ reaches its minimum and at this time,

from (22), we can obtain

$$\mathbf{q} = \mathbf{p}.$$

Therefore, the theorem has been proved, viz. $\mathbf{p} = \arg \min_{\mathbf{q}} d_l(\mathbf{q})$. \square

From this Theorem, if one can obtain an estimation of the cover image, then he (or she) can estimate the embedding ratio for each bit plane of the l LSB planes as follows:

$$\begin{cases} \hat{\mathbf{p}} = \arg \min_{\mathbf{q}} \hat{d}_l(\mathbf{q}), \\ \hat{d}_l(\mathbf{q}) = \sum_{i=1}^n (s_i^{\mathbf{q},l} - \hat{x}_i)^2, \end{cases} \quad (23)$$

where \hat{x}_i denotes the estimation of the cover pixel x_i .

Let the partial derivative of $\hat{d}_l(\mathbf{q})$ with respect to q_k be equal to zero; from the definition of WS in (4) and the definition of $\hat{d}_l(\mathbf{q})$ in (23), one can obtain

$$\begin{aligned} \frac{\partial \hat{d}_l(\mathbf{q})}{\partial q_k} &= 2 \sum_{i=1}^n \left\{ \frac{1}{2} [2^{k-1} - 2(s_i \oplus 2^{k-1})] \right. \\ &\quad \left. \left\{ s_i - \hat{x}_i + \sum_{j=1}^l \frac{q_j}{2} [2^{j-1} - 2(s_i \oplus 2^{j-1})] \right\} \right\} \\ &= \sum_{i=1}^n \{ [2^{k-1} - 2(s_i \oplus 2^{k-1})] (s_i - \hat{x}_i) \} \\ &\quad + \sum_{j=1}^l \sum_{i=1}^n \left\{ \frac{q_j}{2} [2^{k-1} - 2(s_i \oplus 2^{k-1})] \right. \\ &\quad \left. \times [2^{j-1} - 2(s_i \oplus 2^{j-1})] \right\} = 0. \end{aligned} \quad (24)$$

Let

$$\begin{cases} c_{k,0} = \sum_{i=1}^n \{ [2^{k-1} - 2(s_i \oplus 2^{k-1})] (s_i - \hat{x}_i) \} \\ c_{k,j} = \frac{1}{2} \sum_{i=1}^n \{ [2^{k-1} - 2(s_i \oplus 2^{k-1})] \\ \quad \times [2^{j-1} - 2(s_i \oplus 2^{j-1})] \} \end{cases} \quad (25)$$

for $1 \leq j, k \leq l$. Then, the Equation (24) can be written as

$$c_{k,0} + \sum_{j=1}^l q_j c_{k,j} = 0. \quad (26)$$

Let the partial derivatives of $\hat{d}_l(\mathbf{q})$ with respect to q_1, q_2, \dots, q_l be all equal to zero; one can obtain

$$\begin{cases} \frac{\partial \hat{d}_l(\mathbf{q})}{\partial q_1} = c_{1,0} + \sum_{j=1}^l q_j c_{1,j} = 0, \\ \frac{\partial \hat{d}_l(\mathbf{q})}{\partial q_2} = c_{2,0} + \sum_{j=1}^l q_j c_{2,j} = 0, \\ \vdots \\ \frac{\partial \hat{d}_l(\mathbf{q})}{\partial q_l} = c_{l,0} + \sum_{j=1}^l q_j c_{l,j} = 0. \end{cases} \quad (27)$$

Let D denote the determinant of the coefficients of Equation (27), viz.

$$D = \begin{vmatrix} c_{1,1} & c_{1,2} & \cdots & c_{1,l} \\ c_{2,1} & c_{2,2} & \cdots & c_{2,l} \\ \vdots & \vdots & \ddots & \vdots \\ c_{l,1} & c_{l,2} & \cdots & c_{l,l} \end{vmatrix}, \quad (28)$$

and D_k denote the determinant constructed by replacing the elements in the k th column of determinant D with the elements $-c_{1,0}, -c_{2,0}, \dots$, and $-c_{l,0}$, viz.

$$D_k = \begin{vmatrix} c_{1,1} & c_{1,2} & \cdots & c_{1,k-1} & -c_{1,0} & c_{1,k+1} & \cdots & c_{1,l} \\ c_{2,1} & c_{2,2} & \cdots & c_{2,k-1} & -c_{2,0} & c_{2,k+1} & \cdots & c_{2,l} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{l,1} & c_{l,2} & \cdots & c_{l,k-1} & -c_{l,0} & c_{l,k+1} & \cdots & c_{l,l} \end{vmatrix}. \quad (29)$$

Then, from the Cramer's rule, when the determinant D is not equal to zero, one can resolve the Equation (27) to estimate the embedding ratio for each bit plane of l LSB planes as follows:

$$\begin{cases} \hat{p}_1 = q_1 = D_1/D \\ \hat{p}_2 = q_2 = D_2/D \\ \vdots \\ \hat{p}_l = q_l = D_l/D \end{cases} \quad (30)$$

3.3. WS steganalysis algorithm for ID-MLSB steganography

Given a stego-image, one can get access to only the stego-image, but not to the cover image. Therefore, similar to [12], a predictor based on four adjacent pixels is used to estimate the cover image from the given image as follows:

$$\hat{x}_i = (s_{j-1,k} + s_{j,k-1} + s_{j,k+1} + s_{j+1,k})/4, \quad (31)$$

where j and k denote that the estimated pixel \hat{x}_i and the stego pixel s_i are in the j th row and k th column, viz. $i = (j-1) \times N + k$. Then the quantitative steganalysis algorithm of message hidden into each bit plane based on WS can be described as follows.

Input: A given spatial domain image for steganalysis, and the number of LSB planes possibly containing message, *viz.* l .

Output: The embedding ratio for each bit plane of l LSB planes.

Step 1. For each pixel s_i in the given image, do Steps 1.1 and 1.2 to calculate the statistics $c_{k,0}$ and $c_{k,j}$ for $1 \leq j, k \leq l$.

Step 1.1. Estimate the cover pixel \hat{x}_i from formula (31). And when the pixel is in the boundary, estimate the cover pixel by averaging the existing pixels in $\{s_{j-1,k}, s_{j,k-1}, s_{j,k+1}, s_{j+1,k}\}$.

Step 1.2. For $1 \leq j, k \leq l$, compute the increments of $c_{k,0}$ and $c_{k,j}$ from the formula $[2^{k-1} - 2(s_i \oplus 2^{k-1})](s_i - \hat{x}_i)$ and $\frac{1}{2}[2^{k-1} - 2(s_i \oplus 2^{k-1})][2^{j-1} - 2(s_i \oplus 2^{j-1})]$, respectively, and add them to the statistics $c_{k,0}$ and $c_{k,j}$, respectively.

Step 2. Apply the statistics $c_{k,j}$ to formula (28) to compute the determinant D , and apply $c_{k,0}$ and $c_{k,j}$ to formula (29) to compute the determinant D_k , where $1 \leq j, k \leq l$.

Step 3. Apply the determinants D and D_k to the Equation (30) to estimate the embedding ratio \hat{p}_k in the k th LSB plane.

4. EXPERIMENTAL RESULTS AND ANALYSIS

First, an experiment was carried out on an 8-bit uncompressed bitmap image ‘lena.bmp’ (seeing Fig. 3a) with a size of 512×512 pixels to verify the theorem in Section 3.2. Figure 3b shows the squared Euclidean distance between the cover image ‘lena.bmp’ and the weight stego-image with different weights q_1 and q_2 when the ratios of message bits embedded into the

LSB plane and the 2nd LSB plane are $p_1 = 0.3$ and $p_2 = 0.4$, respectively. From this figure, it can be seen that when the weights q_1 and q_2 are equal to the embedding ratios $p_1 = 0.3$ and $p_2 = 0.4$, respectively, the squared Euclidean distance is minimal. This is in accordance with the theorem in Section 3.2.

For evaluating the performance of the proposed steganalysis method, it will be compared with the ID3SPA method in [25] for the case of $l = 3$, *viz.* ID-3LSB steganography. And for ID-3LSB steganography, the proposed steganalysis method is called as ID3WS, where ‘WS’ indicates that the new method is proposed based on WS.

The experimental setup is as follows: 1000 originally very high-resolution color images in ‘tiff’ format were downloaded from <http://photogallery.nrcs.usda.gov>; and then converted to grayscale images in ‘bmp’ format; about a third of them were cropped to leave about 512×512 pixels, another third of them were cropped to leave about 768×768 pixels, and the residual images were cropped to leave about 1024×1024 pixels. (The tool used was Advanced Batch Converter 3.8.20.) The pseudo-random messages were embedded into LSB plane, 2nd LSB plane and 3rd LSB plane of the obtained cover images with the embedding ratios $p_1 \in \{0, 0.1, 0.2, \dots, 1.0\}$, $p_2 \in \{0, 0.1, 0.2, \dots, 1.0\}$ and $p_3 \in \{0.2, 0.4, 0.6, 0.8, 1.0\}$, respectively. Then, the ID3WS and ID3SPA were used to estimate the embedding ratios in the LSB plane, 2nd LSB plane and 3rd LSB plane.

In [27], Böhme *et al.* pointed out that when the quantitative steganalysis methods are used to estimate the embedding ratios or modification ratios, the distribution of the errors is usually heavy-tailed; so the median and interquartile range (IQR) are more representative than the familiar mean and standard deviation in evaluating the estimation error. Therefore, the median and IQR are adopted to evaluate the estimation errors of the proposed ID3WS method and the ID3SPA method.

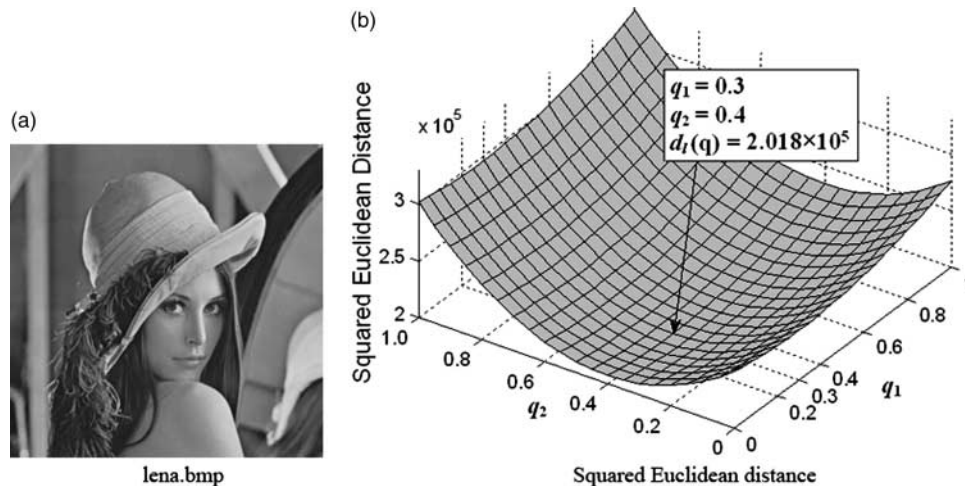


FIGURE 3. The squared Euclidean distance between the cover image ‘lena.bmp’ and the weight stego-image with different weights q_1 and q_2 when the ratios of message bits embedded into the two LSB planes are $p_1 = 0.3$ and $p_2 = 0.4$, respectively. The squared Euclidean distance is minimum in the point explained by the text box.

The median is the 50th percentile of the estimated errors, and the IQR is the difference between the 75th and the 25th percentiles of the estimated errors. For the reason of space, this section does not give all experimental results, and gives only the experimental results for estimating the embedding ratio in the LSB plane when $p_1 = 0, 0.1, 0.2, \dots, 1.0$, $p_2 = 0.4, 0.6, 0.8$ and $p_3 = 0.4, 0.6, 0.8$, and the experimental results for estimating the embedding ratio in the 2nd LSB plane when $p_2 = 0.2, 0.4, 0.6, 0.8, 1.0$ and $p_3 = 0.4, 0.6, 0.8$.

For the embedding ratio in the LSB plane, Figs 4–6 give the estimation errors of the ID3WS and ID3SPA method when the embedding ratio in the 2nd LSB plane is $p_2 = 0.8, 0.6, 0.4$, and the embedding ratio in the 3rd LSB plane is $p_3 = 0.8, 0.6, 0.4$. It can be observed that with the increase of the embedding ratios, the median and IQR of ID3WS's

estimation errors are closer to zero, and the ID3WS method performs better than the ID3SPA method significantly when the embedding ratio in any bit plane is middling or large. Especially when the embedding ratio in any bit plane is close to 1, viz. fully embedding, the typical structural steganalysis method almost fails to estimate the embedding ratios, but contrarily, the ID3WS method can estimate the embedding ratio with a smaller error than that of other cases. And the ID3WS method performs much more stably than ID3SPA method over different embedding ratios. Additionally, the experimental results in Table 1 show the similar conclusion for the estimation error of embedding ratio in the 2nd LSB plane.

The inferiority of the ID3SPA method should recur to the adopted ill-conditioned linear system, which amplifies the error of the assumption based on the ID3SPA method. This is the

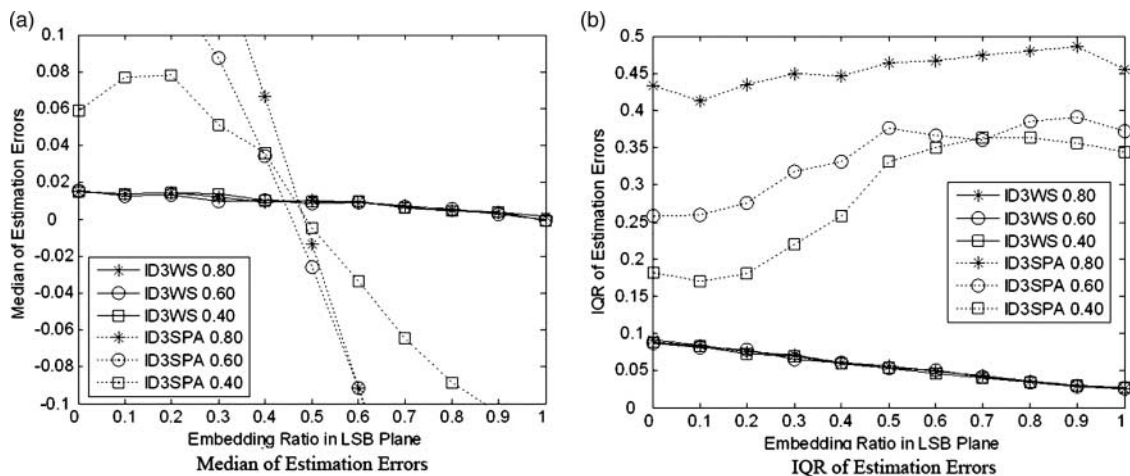


FIGURE 4. Estimation error of embedding ratio in LSB for ID3-LSB steganography when $p_2 = 0.8, 0.6, 0.4$ and $p_3 = 0.8$. (a) Median of estimation errors. (b) IQR estimation errors.

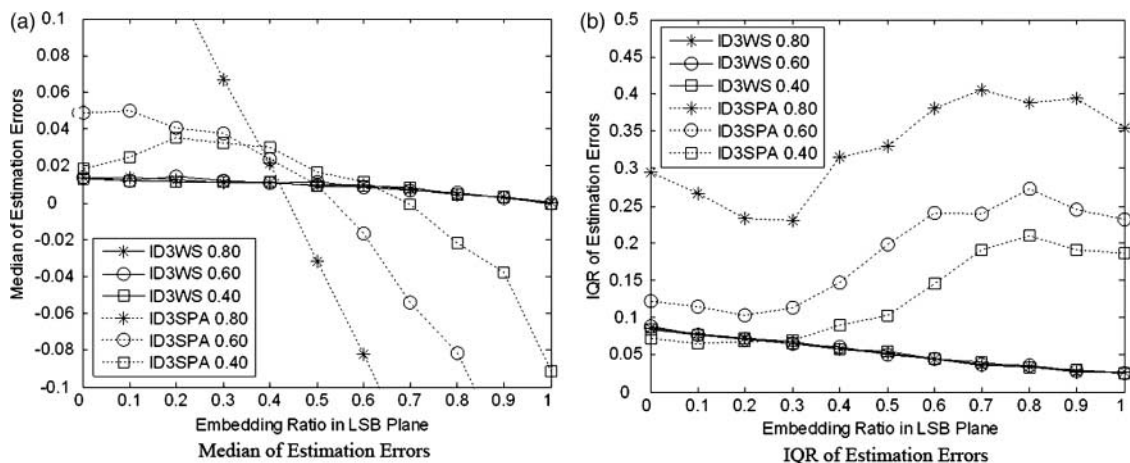


FIGURE 5. Estimation error of embedding ratio in LSB for ID3-LSB steganography when $p_2 = 0.8, 0.6, 0.4$ and $p_3 = 0.6$. (a) Median of estimation errors. (b) IQR estimation errors.

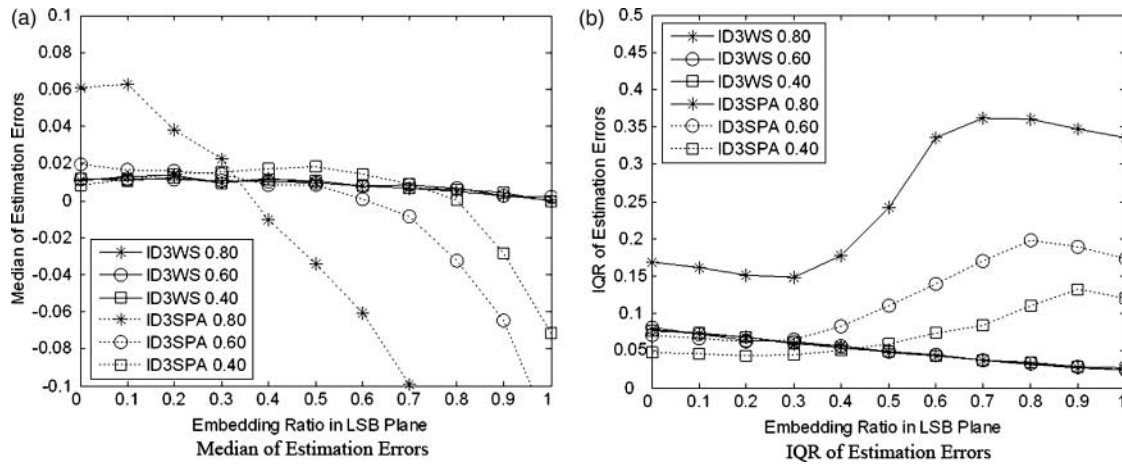


FIGURE 6. Estimation error of embedding ratio in LSB for ID3-LSB steganography when $p_2 = 0.8, 0.6, 0.4$ and $p_3 = 0.4$. (a) Median of estimation errors. (b) IQR estimation errors.

TABLE 1. Estimation error of embedding ratio in the 2nd LSB for ID-3LSB steganography when $p_2 = 0.2, 0.4, 0.6, 0.8, 1.0$ and $p_3 = 0.4, 0.6, 0.8$.

p_3	Metric	Method	p_2				
			0.2	0.4	0.6	0.8	1.0
0.8	Median (10^{-2})	ID3WS	0.95	0.93	0.72	0.47	0.04
		ID3SPA	-1.67	-3.91	-5.95	-8.50	-17.3
	IQR (10^{-2})	ID3WS	4.56	3.96	2.93	2.08	1.40
		ID3SPA	7.48	14.2	25.2	37.3	45.0
0.6	Median (10^{-2})	ID3WS	0.87	0.85	0.80	0.57	-0.02
		ID3SPA	-0.31	-1.06	-1.63	-2.32	-8.16
	IQR (10^{-2})	ID3WS	4.42	3.73	2.83	1.93	1.28
		ID3SPA	3.09	4.37	7.52	13.4	19.3
0.4	Median (10^{-2})	ID3WS	0.85	0.82	0.67	0.50	-0.02
		ID3SPA	0.07	-0.18	-0.66	-1.34	-4.70
	IQR (10^{-2})	ID3WS	4.12	3.50	2.87	1.97	1.22
		ID3SPA	2.32	2.43	2.98	4.75	9.42

The bold data indicate that in the corresponding metrics, the performance of the ID3WS method is better than that of the ID3SPA method.

inherent defect of the structural steganalysis. The performances demonstrate that the proposed methods can fetch up the defect of the structural steganalysis.

However, when the embedding ratios in all the MLSB planes are small, the estimated embedding ratios of the ID3SPA method will own smaller biases and variances. Therefore, in practice, one can use the proposed more stably ID3WS method to estimate the embedding ratios first, and then for the estimated small embedding ratios, use the ID3SPA method to further improve the accuracy.

5. CONCLUSIONS

For ID-MLSb steganography, a new WS model is constructed, and then a new steganalysis method for ID-MLSb steganography is proposed based on the WS model. The experimental results show that the new steganalysis method performs more stably than typical structural steganalysis method with the change of embedding ratios, and outperforms the structural steganalysis on estimation accuracy when the embedding ratio in any bit plane is middling or large. Especially, for the case of embedding ratio in any bit plane is close to 1, the typical structural steganalysis method almost fails to estimate the embedding ratios, but contrarily, the proposed method can estimate the embedding ratio with the highest accuracy. This can effectively fetch up the defects of the structural steganalysis.

In the next steps, we will investigate how to improve the performance of estimating the small embedding ratios, and try to apply the idea of WS to the steganalysis of some adaptive steganography methods.

FUNDING

This work is supported by the National Natural Science Foundation of China (Grant Nos. 60970141, 60902102, 61170032), the Fund of Innovation Scientists and Technicians Outstanding Talents of Henan Province (Grant No. 094200510008) and the Doctoral Dissertation Innovation Fund of Information Science and Technology Institute (Grant No. BSLWCX201002).

REFERENCES

- [1] Fridrich, J. and Golijan, M. (2002) Practical Steganalysis of Digital Images—State of the Art. *Proc. SPIE, Security and Watermarking of Multimedia Contents IV, SPIE 4675*, San Jose, CA, USA, January 21, pp. 1–13. SPIE, Bellingham, WA, USA.

- [2] Lian, S., Kanellopoulos, D. and Ruffo, G. (2009) Recent Advances in Multimedia Information System Security. Technical Report 1. Informatica, Slovenian.
- [3] Lian, S. and Zhang, Y. (2009) *Handbook of Research on Secure Multimedia Distribution*. IGI Global (formerly Idea Group, Inc), Hershey, PA, USA.
- [4] Hosmer, C. and Hyde, C. (2003) Discovering Covert Digital Evidence. *Proc. Workshop on Digital Forensic Research (DFRWS 2003)*, Cleveland, OH, USA, August 6–8. IEEE.
- [5] Fridrich, J., Golijan, M., Hoge, D. and Soukal, D. (2003) Quantitative steganalysis of digital images: estimating the secret message length. *ACM Multimedia Syst. J., Spec. Issue Multimedia Secur.*, **9**, 288–302.
- [6] Fridrich, J., Golijan, M. and Soukal, D. (2004) Searching for the Stego-Key. *Proc. SPIE, Security, Steganography, and Watermarking of Multimedia Contents VI, SPIE 5306*, San Jose, CA, USA, January 18–22, pp. 70–82. SPIE, Bellingham, WA, USA.
- [7] Fridrich, J., Golijan, M., Soukal, D. and Holtyak, T. (2005) Forensic Steganalysis: Determining the Stego Key in Spatial Domain Steganography. *Proc. SPIE, Security, Steganography, and Watermarking of Multimedia Contents VII, SPIE 5681*, San Jose, CA, USA, January 17–21, pp. 631–642. SPIE, Bellingham, WA, USA.
- [8] Ker, A.D. (2008) Locating Steganographic Payload via WS Residual. *Proc. Workshop on Multimedia and Security*, Oxford, UK, September 22–23, pp. 27–32. ACM, New York, NY, USA.
- [9] Fridrich, J., Golijan, M. and Du, R. (2001) Reliable Detection of LSB Steganography in Color and Grayscale Images. *Proc. Workshop on Multimedia and Security*, Ottawa, ON, Canada, October 5, pp. 27–30. ACM, New York, NY, USA.
- [10] Zhang, T. and Ping, X. (2003) A new approach to reliable detection of LSB steganography in natural images. *Signal Process.*, **83**, 2085–2093.
- [11] Dumitrescu, S., Wu, X. and Wang, Z. (2003) Detection of LSB steganography via sample pair analysis. *IEEE Trans. Signal Process.*, **51**, 1995–2007.
- [12] Fridrich, J. and Golijan, M. (2004) On Estimation of Secret Message Length in LSB Steganography in Spatial Domain. *Proc. SPIE, Security, Steganography, and Watermarking of Multimedia Contents VI, SPIE 5306*, San Jose, CA, USA, January 18–22, pp. 23–34. SPIE, Bellingham, WA, USA.
- [13] Ker, A.D. (2004) Improved Detection of LSB Steganography in Grayscale Images. *Proc. 6th Workshop on Information Hiding*, Toronto, Canada, May 23–25, Lecture Notes in Computer Science 3200, pp. 97–115. Springer, Berlin.
- [14] Lu, P., Tang, X.L.Q. and Shen, L. (2004) Improved Detection of LSB Steganography in Grayscale Images. *Proc. 6th Workshop on Information Hiding*, Toronto, Canada, May 23–25, Lecture Notes in Computer Science 3200, pp. 116–127. Springer, Berlin.
- [15] Ker, A.D. and Böhme, R. (2008) Revisiting Weighted Stego-Image Steganalysis. *Proc. SPIE, Security, Forensics, Steganography and Watermarking of Multimedia Contents X, SPIE 6819*, San Jose, CA, USA, January 28–30, pp. 05.1–05.17. SPIE, Bellingham, WA, USA.
- [16] Ker, A.D. (2007) Steganalysis of embedding in two least-significant bits. *IEEE Trans. Inf. Forensics Sec.*, **2**, 46–54.
- [17] Wu, N.-I. and Hwang, M.-S. (2007) Data hiding: current status and key issues. *Int. J. Netw. Secur.*, **4**, 1–9.
- [18] Nguyen, B.C., Yoon, S.M. and Lee, H.-K. (2006) Multi Bit Plane Image Steganography. *Proc. International Workshop on Digital Watermarking*, Jeju Island, Korea, November 8–10, Lecture Notes in Computer Science 4283, pp. 61–70. Springer, Berlin.
- [19] Luo, X., Yang, C., Wang, D. and Liu, F. (2007) LSB steganalysis based on quartic equation. *Lect. Notes Comput. Sci. Trans. Data Hiding Multimedia Secur. II*, **4499**, 68–90.
- [20] Yu, X. and Babaguchi, N. (2008) A Fast and Effective Method to Detect Multiple Least Significant Bits Steganography. *Proc. ACM Symposium on Applied Computing*, Ceará, Brazil, March 16–20, pp. 1443–1447. ACM, New York, NY, USA.
- [21] Yang, C., Liu, F., Luo, X. and Liu, B. (2008) Steganalysis frameworks of embedding in multiple least-significant bits. *IEEE Trans. Inf. Forensics Sec.*, **3**, 662–672.
- [22] Yu, X., Tan, T. and Wang, Y. (2005) Extended Optimization Method of LSB Steganalysis. *Proc. IEEE International Conference on Image Processing*, Genoa, Italy, September 11–14, pp. 1102–1105. IEEE.
- [23] Yu, X. and Babaguchi, N. (2008) Weighted Stego Image based Steganalysis in Multiple Least Significant Bits. *Proc. International Conference on Multimedia and Expo*, Hannover, Germany, June 23–26, pp. 265–268. IEEE.
- [24] Ye, M., Liu, F., Yang, C. and He, X. (2009) Steganalysis Based on Weighted Stego-Image for 2LSB Replacement Steganography. *Proc. 5th International Conf. Intelligent Information and Multimedia Signal Processing*, Kyoto, Japan, September 12–14, pp. 945–948. IEEE.
- [25] Yang, C., Luo, X. and Liu, F. (2009) Embedding Ratio Estimating for Each Bit Plane of Image. *Proc. International Workshop on Information Hiding*, Darmstadt, Germany, June 8–10, Lecture Notes in Computer Science 5806, pp. 59–72. Springer, Berlin.
- [26] Luo, X., Liu, F., Yang, C. and Lian, S. (2010) Modification ratio estimation for a category of adaptive steganography. *Sci. China Ser. F: Inf. Sci.*, **53**, 2472–2484.
- [27] Böhme, R. and Ker, A.D. (2006) A Two-Factor Error Model for Quantitative Steganalysis. *Proc. SPIE, Security, Steganography and Watermarking of Multimedia Contents VIII, SPIE 6072*, San Jose, CA, USA, January 16–19, pp. 59–74. SPIE, Bellingham, WA, USA.