SCIENCE CHINA
Information Sciences

• LETTER •

November 2018, Vol. 61 119103:1–119103:3
https://doi.org/10.1007/s11432-017-9328-2

Extracting hidden messages of MLSB steganography based on optimal stego subset

Chunfang YANG[1,2], Xiangyang LUO[1,2*], Jicang LU[1,2*] & Fenlin LIU[1,2]

[1]*Zhengzhou Information Science and Technology Institute, Zhengzhou* 450001*, China;*
[2]*State Key Laboratory of Mathematical Engineering and Advanced Computing, Zhengzhou* 450001*, China*

Received 11 November 2017/Accepted 29 December 2017/Published online 12 September 2018

Citation Yang C F, Luo X Y, Lu J C, et al. Extracting hidden messages of MLSB steganography based on optimal stego subset. Sci China Inf Sci, 2018, 61(11): 119103, https://doi.org/10.1007/s11432-017-9328-2

Dear editor,

In recent years, steganalysis researchers have tried their best to extract the hidden messages. For example, under the condition of a known embedding position generator, Liu et al. [1] proposed a collision attack algorithm to recover the stego key of least significant bit (LSB) steganography. Fridrich et al. [2] proposed a chi-squared-test-based method to recover the stego key of LSB steganography for the case of an unknown carrier. Under the condition of multiple stego images embedded into the same positions, Ker [3] first proposed locating the payload of the LSB replacement by averaging the weighted stego-image residuals in the same positions of multiple stego images. Subsequently, Ker and Lubenko [4] also proposed a payload location algorithm for LSB matching based on wavelet absolute moments (WAM), which transforms the residuals in the wavelet domain to spatial residuals, and then locates the stego positions by averaging the absolute spatial residuals over multiple stego images. Quach [5,6] adopted the maximum a posteriori method and Markov random fields approach to estimate the cover image, and located the stego position of LSB steganography by combining the differences between multiple stego images and the corresponding estimated cover images. Liu et al. [7] estimated the cover image by compressing the stego image, which had suffered from JPEG compression before embedding the message into the LSBs, and located the payload with higher accuracy. The abovementioned algorithms can be used to effectively extract all hidden message bits of LSB steganography under certain specific conditions. However, there is still a lack of effective algorithms to extract the hidden messages of most steganography algorithms.

This work focuses on an extended LSB steganography—multiple least significant bits (MLSB) steganography, proves the optimal stego subset property of MLSB steganography, and then proposes a stego key recovery algorithm and a payload location algorithm for MLSB steganography.

Let $X = \{x_i\}_{i=1}^N$ be the cover image, $S = \{s_i\}_{i=1}^N$ be the corresponding stego image of MLSB steganography, N denote the number of pixels in the cover or stego images, b denote the number of bits used to store a pixel, l denote the number of bit planes containing the embedded message, and p denote the ratio of stego pixels in the stego image.

MLSB steganography selects certain pixels from a cover image and replaces their l least significant bits with l secret message bits. When the embedded message bits are pseudorandom, for each pixel s_i in the stego image S, the probability that each bit in the l least significant bits of it has been flipped should be $p/2$. Therefore, one can estimate the average cover image by changing the pixel to the opposite direction with the same probability.

* Corresponding author (email: xiangyangluo@126.com, lujicang@sina.com)

© Science China Press and Springer-Verlag GmbH Germany, part of Springer Nature 2018

However, because the embedding ratio is unknown to the steganalyzer, one uses a weight q $(0 \leqslant q \leqslant 1)$ to denote the possible embedding ratio, and obtains the following estimated average cover pixel, which is referred to as the weighted stego pixel:

$$s_i^{(q)} = s_i + \left(2^l - 1 - 2(s_i \bmod 2^l)\right) q/2. \tag{1}$$

Then, the following relationship exists between the weighted stego pixel and the corresponding cover pixel.

Lemma 1. If the pixel s_i contains secret message bits, it follows that

$$\begin{aligned} \mathrm{E}\left\{\left(s_i^{(1)} - x_i\right)^2\right\} &\leqslant \mathrm{E}\left\{\left(s_i^{(q)} - x_i\right)^2\right\} \\ &\leqslant \mathrm{E}\left\{(s_i - x_i)^2\right\}, \end{aligned} \tag{2}$$

where $\mathrm{E}\{\cdot\}$ is the expectation of $\cdot$, $0 \leqslant q \leqslant 1$; if pixel s_i does not contain secret message bits, it follows that

$$\begin{aligned} \mathrm{E}\left\{\left(s_i^{(1)} - x_i\right)^2\right\} &\geqslant \mathrm{E}\left\{\left(s_i^{(q)} - x_i\right)^2\right\} \\ &\geqslant \mathrm{E}\left\{(s_i - x_i)^2\right\}. \end{aligned} \tag{3}$$

The equal signs on the left sides of (2) and (3) hold if and only if the weight $q = 1$, and the equal signs on the right sides of (2) and (3) hold if and only if the weight $q = 0$.

A detailed proof of Lemma 1 is presented in Appendix A.

One can divide the pixels in the stego image S into the suspected stego pixel subset G and nonstego pixel subset $H = S - G$, and construct the weighted stego pixel subsets $G^{(1)} = \{s_i^{(1)} | s_i \in G\}$ and $H^{(0)} = \{s_i^{(0)} | s_i \in H\}$. Theorem 1 can then be derived.

Theorem 1. If one constructs a specific weighted stego pixel set $S^{(G)}$ by uniting the weighted stego pixel subsets $G^{(1)}$ and $H^{(0)}$, namely $S^{(G)} = G^{(1)} \bigcup H^{(0)}$,

$$\begin{aligned} U = \arg\min_G \mathrm{E}\Bigg\{ &\sum_{s_i \in G} \left(s_i^{(1)} - x_i\right)^2 \\ &+ \sum_{s_i \in S-G} (s_i - x_i)^2 \Bigg\}, \end{aligned} \tag{4}$$

where U is the stego pixel subset in the stego image S. The subset G that minimizes the expectation of the squared Euclidean distance between $S^{(G)}$ and the cover pixel set is referred to as the optimal stego subset.

A detailed proof of Theorem 1 is presented in Appendix B.

Therefore, if one can obtain an estimated cover image $\hat{X} = \{\hat{x}_i\}_{i=1}^N$, it is possible to estimate the stego pixel subset as follows:

$$\begin{aligned} \hat{U} &= \arg\min_G \Bigg[\sum_{s_i \in G} \left(s_i^{(1)} - \hat{x}_i\right)^2 \\ &\quad + \sum_{s_i \in S-G} \left(s_i - \hat{x}_i\right)^2 \Bigg] \\ &= \arg\min_G \sum_{s_i \in G} \hat{r}_i, \end{aligned} \tag{5}$$

where

$$\begin{aligned} \hat{r}_i = &(s_i - \hat{x}_i)(2^l - 1 - 2(s_i \bmod 2^l)) \\ &+ (2^l - 1 - 2(s_i \bmod 2^l))^2/4. \end{aligned} \tag{6}$$

Then, one can extract the hidden messages of MLSB steganography from the estimated stego pixels. However, the number of all possible subsets of pixels in a stego image is 2^N, which is too large to complete the search within an acceptable time without any priori knowledge. Moreover, stego pixels that are unchanged during message embedding are easily confused with the nonstego pixels. Thus, more knowledge about the implementation of MLSB steganography is required in order to efficiently extract the hidden message.

From above idea, when one knows the embedding position generator that should be fed a stego key, but does not know the stego key, the number of possible subsets is equivalent to the cardinality of the stego key space, which may be significantly smaller than 2^N, and the elements in the optimal stego subset can be selected by the correct stego key. Therefore, one can recover the stego key by searching for the correct stego key from the stego key space to find the subset with the minimum sum of $\hat{r}_i$ as follows (a detailed algorithm is described in Appendix C):

$$\hat{K} = \arg\min_{t \in \Omega} \sum_{i \in \Gamma(t,L),} \hat{r}_i, \tag{7}$$

where $\hat{K}$ denotes the recovered stego key, $\Omega = \{k_0, k_1, \ldots, k_D\}$ denotes the stego key space, D denotes the cardinality of the stego key space, and $\Gamma(t, L)$ denotes the set of L positions generated with the possible key t. Then, the hidden message can be extracted with the recovered stego key.

When one does not know the embedding position generator, but possesses multiple stego images embedded in the same positions, one can locate the stego positions as follows (a detailed algorithm is described in Appendix D):

$$\hat{Z} = \arg\min_W \sum_{i \in W,} \sum_{j=1}^{M} \hat{r}_{i,j}, \tag{8}$$

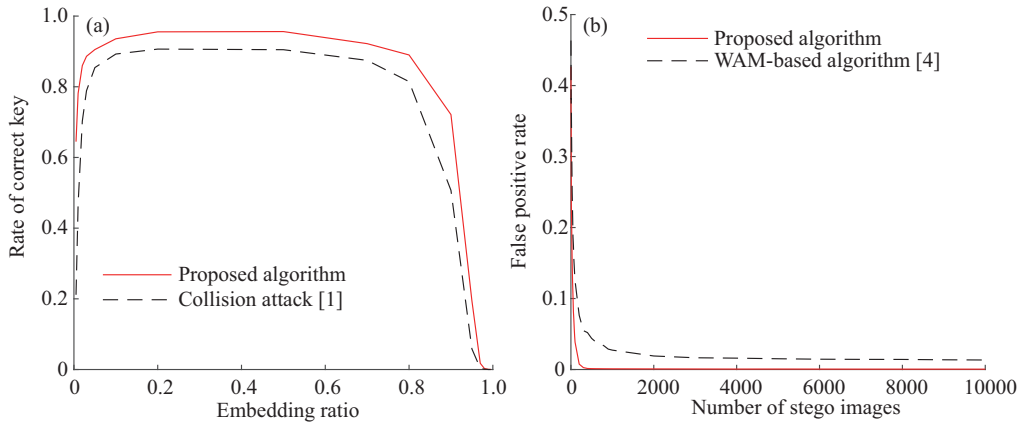

Figure 1 (Color online) Performance of the proposed algorithms and previous algorithms for 2LSB steganography. (a) Performance of the stego key recovery algorithms; (b) performance of the payload location algorithms when the embedding ratio $p = 0.5$.

where $\hat{Z}$ denotes the estimated stego position set, W denotes the selected position subset, M denotes the number of stego images embedded in the same positions, $s_{i,j}$ denotes the i-th pixel in the j-th stego image, $\hat{x}_{i,j}$ denotes the estimated cover pixel, and $\hat{r}_{i,j}$ is computed as follows:

$$\begin{aligned}\hat{r}_{i,j} &= (s_{i,j} - \hat{x}_{i,j})(2^l - 1 - 2(s_{i,j} \bmod 2^l)) \\ &\quad +(2^l - 1 - 2(s_{i,j} \bmod 2^l))^2/4.\end{aligned} \tag{9}$$

Experiments. The proposed algorithms were tested in 10000 gray cover images of size 512 pixels × 512 pixels, randomly cropped from high-resolution uncompressed colored "tiff" images downloaded from http://agents.fel.cvut.cz/stegodata/RAWs/. Figure 1(a) shows that the proposed stego key recovery algorithm can recover the stego keys with significantly higher success rates than the collision attack algorithm when the embedding ratio of 2LSB steganography is smaller than 0.95. Figure 1(b) shows that the proposed payload location algorithm can locate the payload with a lower false positive rate, and the false positive rate quickly approaches 0 with the increasing number of stego images. Further experimental results and analysis are supplied in Appendix E.

Conclusion. This study proved the optimal stego subset property of MLSB steganography, and proposed a stego key recovery algorithm and a payload location algorithm for two special cases. The experimental results showed that the proposed algorithms can recover the stego key and locate the payload with higher accuracy than previous algorithms. However, certain issues remain unaddressed; for example, the algorithms cannot be directly applied to other steganography and other media, such as speech streams [8].

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant Nos. 61302159, 61401512, 61772549, 61379151, 61602508, U1736214) and Science and Technology Research Project of Henan Province, China (Grant No. 152102210005).

Supporting information Appendixes A–E. The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 Liu J F, Tian Y G, Han T, et al. Stego key searching for LSB steganography on JPEG decompressed image. Sci China Inf Sci, 2016, 59: 032105
2 Fridrich J, Goljan M, Soukal D. Searching for the stego-key. In: Proceedings of SPIE 5306, Security, Steganography, and Watermarking of Multimedia Contents VI, San Jose, 2004. 70–82
3 Ker A D. Locating steganographic payload via WS residuals. In: Proceedings of 10th Workshop on Multimedia and Security, Oxford, 2008. 27–31
4 Ker A D, Lubenko I. Feature reduction and payload location with WAM steganalysis. In: Proceedings of SPIE 7254, Media Forensics and Security, San Jose, 2009. 72540A
5 Quach T T. Optimal cover estimation methods and steganographic payload location. IEEE Trans Inform Forensic Secur, 2011, 6: 1214–1222
6 Quach T T. Cover estimation and payload location using Markov random fields. In: Proceedings of SPIE 9028, Electronic Imaging, Media Watermarking, Security, and Forensics, San Francisco, 2014. 90280H
7 Liu J F, Tian Y G, Han T, et al. LSB steganographic payload location for JPEG-decompressed images. Digital Signal Process, 2015, 38: 66–76
8 Huang Y F, Tao H Z, Xiao B, et al. Steganography in low bit-rate speech streams based on quantization index modulation controlled by keys. Sci China Technol Sci, 2017, 60: 1585–1596