2018 IEEE Third International Conference on Data Science in Cyberspace

Ensemble Steganography

Chuan Qin
CAS Key Laboratory of
Electro-magnetic Space Information
University of Science
and Technology of China Hefei, China
Email: qc94@mail.ustc.edu.cn

Wenbo Zhou
CAS Key Laboratory of
Electro-magnetic Space Information
University of Science
and Technology of China Hefei, China
Email: welbeckz@mail.ustc.edu.cn

Weiming Zhang
CAS Key Laboratory of
Electro-magnetic Space Information
University of Science
and Technology of China Hefei, China
Tel.:0551-63600863
Email: zhangwm@ustc.edu.cn

Nenghai Yu
CAS Key Laboratory of
Electro-magnetic Space Information
University of Science
and Technology of China Hefei, China
Tel.:0551-63600681
Email: ynh@ustc.edu.cn

***Abstract*—Currently, selection-channel-aware (SCA) steganalysis is the most successful technique for detecting steganography. It estimates the modification probabilities to improve its performance. How to counter SCA steganalysis is one of the most challenging works in steganography. In this paper, we propose ensemble steganography which introduces random cost functions to disturb the estimation of the modification probabilities of cover elements. The experimental results imply that ensemble steganography improves the performance of HILL and UNIWARD in countering SCA steganalysis and the security of ensemble steganography increases with the number of cost functions and the differences among them.**

I. Introduction

Steganography is a technique of hiding information into ordinary digital media while drawing minimal suspicion [1] [2]. It is challenging to design steganographic algorithms for various cover sources due to the lack of accurate models. Currently, minimizing the distortion between the cover and the corresponding stego object is the most successful approach for designing content adaptive steganography. By assigning a cost to each modified cover element (e.g., pixel in the spatial image) the distortion is obtained, and the messages are embedded with practical codes such as Syndrome-trellis codes (STCs) [3] while minimizing the sum of distortions caused by all the modifications.

HUGO (highly undetectable stego) [4] is the first method based on the framework of minimal distortion steganography. In HUGO, the cost of modifying a pixel is defined as the changes of SPAM (subtractive pixel adjacency matrix) features [5] that the modification causes. And higher costs are applied to those pixels which cause more deviation in the feature vector. This design of cost function makes the modification of HUGO clustered in textural areas. However, steganalyzers with higher dimensions, such as SRM (spatial rich models) [6], can make HUGO be vulnerable.

In SRM, various filters are utilized to generate the predicted residuals, so it overcomes the diversity of cover sources and further exploits the correlations among pixels. Therefore, steganography should elaborate its definition of smooth or textural areas. One pixel should be assigned high cost if it can be predicted in as long as one direction (e.g.in smooth areas and at clean edges) and low if it is unpredictable in every direction (in textural areas). With this insight, Hulob and Fridrich proposed WOW [7] by which the pixels that are more predictable by a bank of directional filters are assigned high costs. By this way, HUGO is improved by WOW under the detection of SRM [6]. UNIWARD (universal wavelet relative distortion) [8] bears similar performance to WOW while simplifies the cost function of WOW and generalizes it to be more suitable for embedding in an arbitrary domain, the spatial domain and DCT domain included. However, WOW and UNIWARD assign some pixels in textural areas, which may be suitable for carrying data, with high costs. So Li et al. proposed the method HILL [9]. In HILL, the low costs of pixels in textural areas are spread to their neighborhood by a low-pass filter, which avoid pixels with high cost values in textural areas and improves WOW.

Apparently, the modification probabilities and the embedding locations are closely related to the cost functions of the state-of-the-art adaptive steganographic methods. Based on this logic, steganalysis follows up. Tang et al. proposed an adaptive steganalytic scheme [10] (called tSRM) for the method WOW, which narrows down the possible embedding areas for a given suspicious image and extract features only from such areas. Because tSRM only targets at WOW, Tomas Denemark et al. proposed selection-channel-aware (SCA) rich model (called maxSRM) [11]. MaxSRM improves tSRM by forming the co-occurrence matrices that the maximum estimated modification probability of neighboring pixels as a weight. The similar idea is applied in meanSRM [12], which

978-1-5386-4210-8/18/$31.00 ©2018 IEEE
DOI 10.1109/DSC.2018.00093

utilize the mean estimated modification probability of a group of pixels as a weight coefficient. Recently, Denemark and Fridrich proposed an improving SCA steganalysis feature [13], which use the residuals in SRM instead of co-occurrence matrices.

The inevitable loophole of the state-of-the-art adaptive steganographic algorithms has been exploited by steganalyzers using various kinds of SCA steganalysis features. However, please note that some minimal-distortion steganographic methods define cost functions in completely different ways, which leads to very different costs and modification probabilities to the same pixel. According to the experimental result, the performance of maxSRM will face a significant decrease when using a different modification probability matrix. So In this paper, we proposed a novel scheme of steganography which randomly assigns the cost of pixels to disturb the estimation of modification probability of maxSRM steganalyzers. We divide the cover image into blocks of the same size. Then a cost function, which is called pre-allocating function, will be utilized to allocating message bits among those blocks. However, the modification locations inside a block is determined by a cost function randomly taken from a group of steganography methods. We call such methods disturbing functions. The scheme described above is called ensemble steganography. According to the experimental results, when suitable parameters are chosen (such as pre-allocating function, disturbing functions and block size), we find this technique can help improve recent additive distortion steganography methods to resist SCA steganalyzers.

The rest of this paper is organized as follows. After introducing notations and the framework of minimal distortion steganography in Section II, in Section III, we propose ensemble steganography which creates random modification probability maps. Results of comparative experiments are elaborated in Section IV to demonstrate the effectiveness of the proposed scheme. Conclusion and future work are given in Section V.

II. Preliminaries

A. Notations

In this paper, matrices, vectors and sets are written in bold face, and the entropy function is denoted $H(\pi_1, ..., \pi_k)$ for $\sum_{i=1}^{k} \pi_i = 1$.

The cover (and the corresponding stego) sequence is denoted by $\boldsymbol{X} = (x_1, x_2, ..., x_n)$, where signal x_i is an integer, such as the gray value of a pixel. The embedding operation on x_i is formulated by the range I_i. An embedding operation is called binary if and ternary if for all i. For example, the ± 1 embedding operation is ternary embedding with $I_i = \{x_i - 1, x_i, x_i + 1\}$.

B. Minimal Additive Distortion Steganography

In the model established in [3], the cover $\boldsymbol{X}$ is assumed to be fixed, so the distortion introduced by changing $\boldsymbol{X}$ to $\boldsymbol{Y} = (y_1, y_2, ..., y_n)$ can be simply denoted as $D(\boldsymbol{X}, \boldsymbol{Y}) = D(\boldsymbol{Y})$. Assume that the embedding algorithm changes $\boldsymbol{X}$ to $\boldsymbol{Y} \in \mathcal{Y}$ with probability $\pi(\boldsymbol{Y}) = P(Y = \boldsymbol{Y})$ which is called the modification probability, and thus the sender can send to up to $H_k(\pi)$ bits of message on average with distortion $E_\pi(D)$ such that

$$H(\pi) = -\sum_{\boldsymbol{Y} \in \mathcal{Y}} \pi(\boldsymbol{Y}) \log \left(\pi(\mathbf{Y})\right), \tag{1}$$

$$E_\pi(D) = \sum_{\boldsymbol{Y} \in \mathcal{Y}} \pi(\boldsymbol{Y}) D(\boldsymbol{Y}). \tag{2}$$

For a given message length L, the sender wants to minimize the average distortion, which can be formulated as the following optimization problems:

$$\min_{\pi} \ E_\pi(D), \tag{3}$$

$$subject\ to\ H(\pi) = L. \tag{4}$$

Following the maximum entropy principle, the optimal has a Gibbs distribution [3]:

$$\pi_\lambda(Y) = \frac{1}{Z(\lambda)} \exp(-D(Y)), \tag{5}$$

where $Z(\lambda)$ is normalizing factor such that

$$\mathrm{Z}(\lambda) = \sum_{\mathbf{Y} \in \mathcal{Y}} \exp(-D(\mathbf{Y})). \tag{6}$$

The scalar parameter $\lambda > 0$ can be determined by the payload constraint (4). In fact, as proven in [14], the entropy in (4) is monotonically decreasing in λ, thus for a given L in the feasible region, λ can be quickly determined by binary search.

In particular, if the embedding operations on x_i's are mutually independent, the distortion introduced by changing $\boldsymbol{X}$ to $\boldsymbol{Y}$ can be thought to be additive, and are measured by $D(\boldsymbol{Y}) = \sum_{i=1}^{n} \rho_i(y_i)$, where $\rho_i(y_i) \in \mathbb{R}^*$ is the cost of changing the ith cover element x_i to $y_i (y_i \in I_i, i = 1, 2, ..., n)$. In this case, the optimal π is given by

$$\pi_i(y_i) \ = \ \frac{exp(-\lambda \rho_i(y_i))}{\sum_{y_i \in I_i} exp(-\lambda \rho_i(y_i))}, \ i = 1, 2, ..., n. \tag{7}$$

When varying $\lambda \in (0, \infty)$, we can derive a relation between $H(\pi)$ and $E_\pi(D)$, which is called the *rate-distortion bound* [14]. Practical coding methods work under this bound.

In this paper, we consider the case of ternary embedding with the range $I_i = \{x_i - 1, x_i, x_i + 1\}$. Usually, we assume that $+1$ and -1 cause the same cost and thus define

$$\rho_i(x_i - 1) = \rho_i(x_i + 1) \triangleq \rho_i \in [0, +\infty). \tag{8}$$

And with Eq.(8), it can be assumed that

$$\begin{cases} \pi_i(x_i - 1) = \pi_i(x_i + 1) \triangleq \tau_i \in \left[0, \frac{1}{3}\right], \\ \pi_i(x_i) = 1 - 2\tau_i. \end{cases} \tag{9}$$

For additive distortion, simulating optimal embedding enables us to test the security of a steganographic method, but once the distortion function is properly defined, we can replace the optimal embedding simulator with some off-the-shelf coding methods such as STCs (Syndrome-Trellis Codes) [3], which can approach the lower rate-distortion bound.

(a) HILL (b) UNIWARD (c) WOW

Fig. 1. The modification probability maps of different cost functions. The brightness in (a) to (c) is scaled to [0,1], where 0 is the darkest (lowest probability) and 1 is the brightest (highest probability).

TABLE I
USE DIFFERENT PROBABILITY MAPS TO EXTRACT FEATURES AND TRAIN CLASSIFIER TO DETECT (PAYLOAD = 0.3BPP).

Probability map	Cost function		
	HILL	UNIWARD	WOW
HILL	.2628 ± .0022	.2931 ± .0033	.2612 ± .0048
UNIWARD	.2671 ± .0020	.2350 ± .0021	.2064 ± .0017
WOW	.2699 ± .0017	.2460 ± .0024	.1896 ± .0027

III. ENSEMBLE STEGANOGRAPHY

A. Motivation

Under the framework of minimal distortion steganography, the major difference among current works is the design of cost functions. According to Eq.(5), different cost assignment will lead to different modification probability map. And Fig.1 shows visually how different the modification probability maps of those cost functions are.

Estimating the modification probabilities is crucially important for maxSRM steganalyzers. So we assume that disturbed estimation of modification probabilities would lead to higher testing error of steganalysis. To verify this assumption, we use three state-of-the-art methods to create stego images while the maxSRM steganalyzer uses these methods to generate estimated modification probability maps. The results are shown in Table I.

B. Proposed Method

According to the experimental results above, disturbing the estimation of modification probability map by using different cost functions can decrease the performance of maxSRM steganalysis. So we propose to create such disturbance by integrating several cost functions together, which are disturbing functions. The proposed scheme is called ensemble steganography. In ensemble steganography, we assume that the cover $\boldsymbol{X}$ consists of n pixels and the message $\boldsymbol{M}$ consists of l bits, and a function set including r cost functions such that $\{f_1, f_2, ..., f_r\}$. The procedure of ensemble steganography be implemented in 4 steps, and the diagram is shown in Fig.2.

1) Divide the pixels of $\boldsymbol{X}$ into s blocks with the same size $q = \frac{n}{s}$. By using the stego-key: k, randomly group these blocks into r groups $\boldsymbol{G}_1, \boldsymbol{G}_2, ..., \boldsymbol{G}_r$, such that each group includes h blocks and the ith group is denoted as $\boldsymbol{G}_i = \{\boldsymbol{B}_{i,1}, ..., \boldsymbol{B}_{i,h}\}$ $(1 \leq i \leq r)$. Herein, $h = \frac{s}{r}$, and without loss of generality, we assume both $q = \frac{n}{s}$ and $h = \frac{s}{r}$ are integers.
2) Select a cost function f^* to allocate message $\boldsymbol{M}_i$ with length l_i to the block group $\boldsymbol{G}_i, (1 \leq i \leq r)$. The cost function f^* is called pre-allocating function. In this step, with a specific cost function f^*, we pre-define costs on all pixels. According to the costs and the message length l, we can calculate the modification probability of each pixel with Eq.(7). The pixels in $\boldsymbol{G}_i$ are denoted as $x_{i,1}, x_{i,2}, ..., x_{i,h}$, and the modification probability on $x_{i,j}(1 \leq j \leq h)$ as $\{\pi_i(x_i - 1), \pi_i(x_i + 1)\}$, where $\pi_i(x_i - 1)$ and $\pi_i(x_i + 1)$ are denoted as τ_i according to Eq.(9). The message length allocated to $\boldsymbol{G}_i$ is

$$l_i = \sum_{j=1}^{h} 2H(\tau_i) + H(1 - 2\tau_i). \tag{10}$$

3) For each group $\boldsymbol{G}_i (1 \leq i \leq r)$, define costs on pixels of each block $\boldsymbol{G}_i$ with the ith cost function f_i which is called disturbing function, and then concatenate these pixel blocks and embed the message $\boldsymbol{M}_i$ into the $\boldsymbol{G}_i$ with STCs, which generate a series of stego blocks. Please note that we pad the cover blocks with their neighboring pixels to keep the correlations of costs defining in block edges when define costs in blocks with f_i.
4) Put back each stego block to the position of the corresponding cover block and generate the stego image $\boldsymbol{Y}$.

Furthermore, in practical scenario, the stego-key k, block size q and group number r are shared between the sender and the receiver beforehand, the same as STCs [3] generator matrix $\boldsymbol{H}$.

C. Pseudo-code Procedure

We provide the pseudo-code that describes implementation of message embedding and extraction in Algorithm 1 and Algorithm 2 respectively.

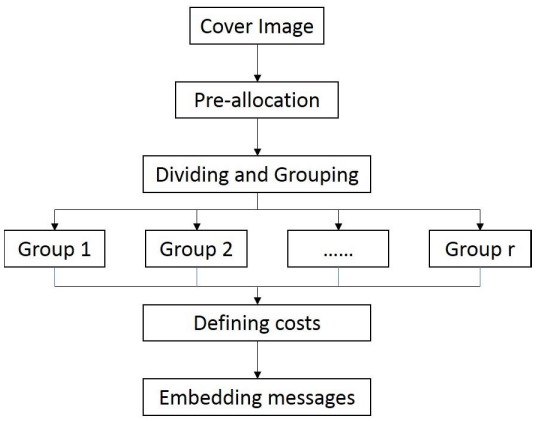


Fig. 2. The diagram of the proposed scheme ensemble steganography

Algorithm 1 Ensemble Steganography Embedding

INPUT: A cover image $\boldsymbol{X}$ with n pixels, l bits of message $\boldsymbol{M}$ determining the payload $\gamma = l / n$, a stego key with value k to shuffle the random sequence of image blocks, the number of groups and disturbing functions r.

OUTPUT: A stego image $\boldsymbol{Y}$.

1: Divide the $\boldsymbol{X}$ into s blocks with the same size $q = \frac{n}{s}$;
2: Randomly group image blocks into r groups $\boldsymbol{G}_1, \boldsymbol{G}_2, ..., \boldsymbol{G}_r$;
3: Utilize cost function f^* to pre-define the cost of all the pixels and allocate the message array $\boldsymbol{M}_i$ to each group $\boldsymbol{G}_i$ with length l_i;
4: Use disturbing function $f_i (1 \leq i \leq r)$ to define costs on pixels of each block in $\boldsymbol{G}_i$ respectively;
5: Embed the message array $\boldsymbol{M}_i$ into the cover group $\boldsymbol{G}_i$ with STCs;
6: Reshape the cover group and put back the stego blocks to generate the stego image Y.

Algorithm 2 Ensemble Steganography Extraction

INPUT: The received stego image $\boldsymbol{Y}$, the stego-key k, the block size q, the number of groups and disturbing functions r.

OUTPUT: The extracted message $\boldsymbol{M}$.

1: Divide the stego image into s blocks of the same size q, and group them into r groups $\boldsymbol{G}_i$ according to the random index array generated by the stego key k;
2: Extract the number of message bits l_i of each group $\boldsymbol{G}_i$, and extract the message $\boldsymbol{M}_i$;
3: Concatenate the message arrays to get the message $\boldsymbol{M}$.

TABLE II
STEGANALYSIS EXPERIMENTAL RESULTS OF DIFFERENT PRE-ALLOCATING FUNCTIONS

Probability map	Pre-allocating function		
	HILL	UM-HILL	UNIWARD
HILL	.2698 ± .0028	.2670 ± .0023	.2714 ± .0036
UM-HILL	.2679 ± .0020	.2661 ± .0039	.2686 ± .0026
UNIWARD	.2576 ± .0023	.2523 ± .0024	.2495 ± .0027

IV. Experimental Results and Discussion

A. Setups

All experiments in this paper are carried out on BOSSbase 1.01 [15] with the amount of 10,000 gray-scale images with size 512×512 pixels. Furthermore, the steganalytic experiments are conducted on a CPU cluster with 12 Xen E5-2650 v4 CPU cards. The detectors are trained as binary classifiers using the FLD. The ensemble classifier by default minimizes the total classification error probability $P_E = min_{P_{FA}} \frac{1}{2} (P_{FA} + P_{MD})$, where P_{FA} and P_{MD} are the false-alarm probability and the missed-detection probability respectively. The ultimate security is qualified by average error rate $\overline{P_E}$ by 10 times random 5000/5000 database splits, and larger $\overline{P_E}$ means stronger security. The detector is trained using the SCA feature maxSRMd2 [11]. We use the optimal embedding simulator as default for all algorithms in steganalysis experiments.

Since the number of cost functions is quite limited, to increase the number and diversity of cost functions, two techniques of fluctuating costs which can be used for any cost functions are introduced below.

1) Microscope Algorithm: Chen et al. proposed a method [16] which aims at designing better cost functions by exposing more details of texture of the cover image. In their methods, a "Microscope" which is unsharp masking (UM) is utilized to highlight the details of the cover image, and the enhanced image is called the auxiliary image on which previous steganography methods are used to define cost. The final cost will be get after the embedding distortion being smoothed by a low-pass filter.

2) Game-theoretic Algorithm: Recently, Li el al. proposed a new idea in [17], in which game theory has been taken into consideration for designing cost function. The main idea of [17] is utilize an existing adaptive steganographic method to define a basic cost function, and then a bias function is defined in the framework of game theory to adjust the distribution of modification probabilities.

The steganographic method adopting the above two schemes would be prefixed with "UM" and "bias" respectively, such as UM-HILL, UM-UNIWARD, bias-HILL and bias-UNIWARD.

B. The Selection of Pre-allocating Function

In this experiment, HILL, UM-HILL and UNIWARD are compared as pre-allocating function. The experiment is set up in this way, at the side of the sender, the disturbing functions are kept the same while the pre-allocating function is changed.

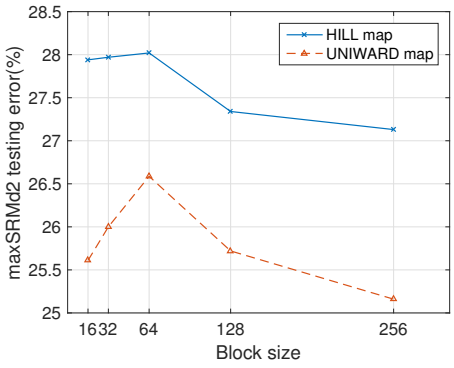


Fig. 3. Testing error of maxSRMd2 with different block sizes.

At the side of the steganalyzer, for a steganographic method with certain pre-allocating function, the three pre-allocating functions will be used to generate the modification probability maps of SCA and extract the features for training the binary classifiers. And the payload is fixed with 0.3 bpp. Finally, we compare the performances of the three pre-allocating functions based on the testing errors of the three steganalyzers described above.

According to the results shown in Table II, we found that HILL is the optimal pre-allocating function of the proposed method.

C. Determination of Block Size

In ensemble steganography, we want to introduce more disturbance for maxSRM steganalyzers, which increases with the number of blocks. However, with the block size decrease, some cost defining rules will be violated. For instance, HILL [9] applies *the cost spreading rule*, in which pixels spread the costs to their neighborhood. However, in ensemble steganography with small block size, most of the neighboring pixels are defined by other cost functions. So there should be a balanced point of these two impacts, so in this section, we set up the experiment to optimize the value of block size. The block sizes being tested are all exact divisions of 512 (the width and height of spatial cover image in BOSSbase 1.01) to avoiding impact of marginal blocks which have different sizes. And without loss of generality, the pre-allocating function we use is HILL and the disturbing functions are HILL and UNIWARD, and both of HILL and UNIWARD are used to generate modification probability maps of maxSRMd2 steganalysis. According to Fig.3, the optimal block size should be 64.

D. The Selection of Disturbing Functions

After fixing the pre-allocating function and block size, the disturbing functions should be considered. Four sets of disturbing functions are tested. 1) HILL and UNIWARD; 2) HILL group: HILL, UM-HILL and bias-HILL; 3) UNIWARD group: UNIWARD, UM-UNIWARD and bias-UNIWARD; 4) mixture group: HILL group and UNIWARD group. Every disturbing function of each set is utilized to generate modification probability maps for maxSRMd2 steganalyzer. And the results are shown in Fig.4 and Fig.5.

According to the results shown in Fig.4 and Fig.5, we found out that the combination HILL and UNIWARD has the best overall performance, and mixture group is better than HILL group and UNIWARD group. Please note that HILL and UNIWARD have quite different modification probability maps as the experimental results in Table I indicates. It provides the reason why HILL and UNIWARD combination has the best performance. Furthermore, by comparing the performance of the mixture group with HILL group and UNIWARD group we can draw a conclusion that the security of ensemble steganography increases with the number of disturbing functions.

V. Conclusion

In this paper, we proposed ensemble steganography to counter SCA steganalysis. According to the experimental results, we found ensemble steganography can improve the ability of most previous works in resisting SCA steganalysis. Furthermore, with the increase of disturbing methods and differences among disturbing functions, ensemble steganography would have better steganalytic performances. So we believe ensemble steganography is a valid strategy to counter SCA steganalysis.

In the present paper, we only utilize two original steganographic methods as disturbance in spatial images with additive distortion for ± 1 embedding. In our future work, applying ensemble steganography to JPEG image and utilizing more steganographic methods which are based on non-additive framework is an interesting direction. Furthermore, to create more significant disturbances, looking for a low-dimensional criterion which can fit the increase of testing error when using disturbed maps is an challenging direction.

Acknowledgment

This work was supported in part by the Natural Science Foundation of China under Grant U1636201 and 61572452.

References

[1] J. Fridrich, *Steganography in digital media: principles, algorithms, and applications*. Cambridge University Press, 2009.

[2] B. Li, J. He, J. Huang, and Y. Q. Shi, "A survey on image steganography and steganalysis," *Journal of Information Hiding and Multimedia Signal Processing*, vol. 2, no. 2, pp. 142–172, 2011.

[3] T. Filler, J. Judas, and J. Fridrich, "Minimizing additive distortion in steganography using syndrome-trellis codes," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 920–935, 2011.

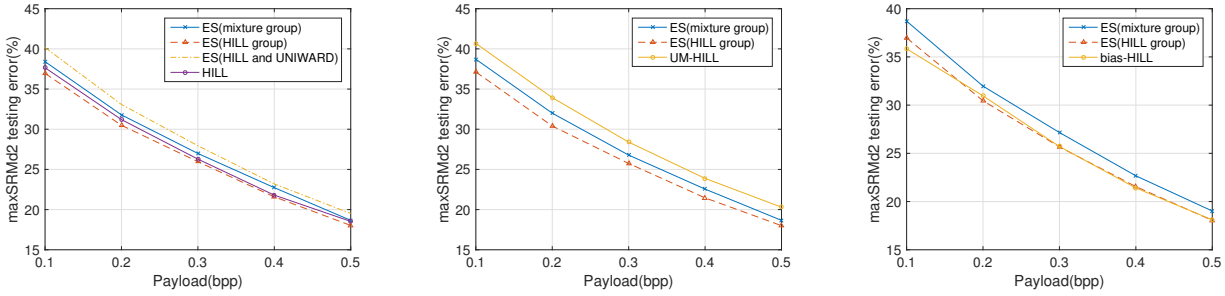
[4] T. Pevnỳ, T. Filler, and P. Bas, "Using high-dimensional image models to perform highly undetectable steganography," in *International Workshop on Information Hiding*. Springer, 2010, pp. 161–177.

[5] T. Pevny, P. Bas, and J. Fridrich, "Steganalysis by subtractive pixel adjacency matrix," *IEEE Transactions on information Forensics and Security*, vol. 5, no. 2, pp. 215–224, 2010.

[6] J. Fridrich and J. Kodovsky, "Rich models for steganalysis of digital images," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 3, pp. 868–882, 2012.

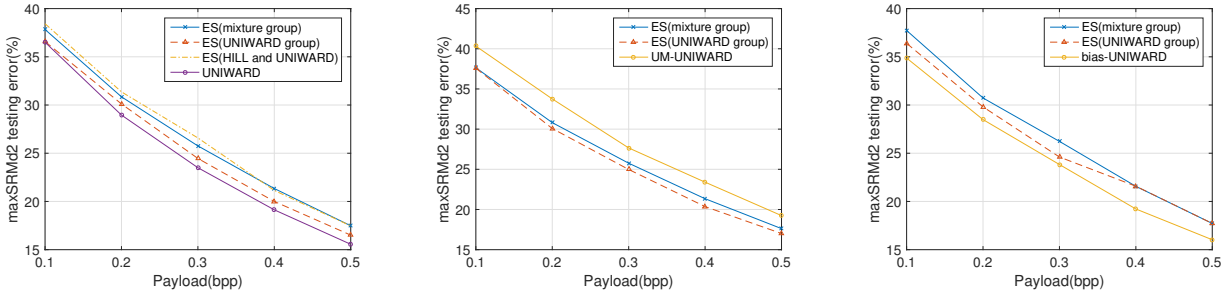
[7] V. Holub and J. Fridrich, "Designing steganographic distortion using directional filters," in *Information Forensics and Security (WIFS), 2012 IEEE International Workshop on*. IEEE, 2012, pp. 234–239.

[8] V. Holub, J. Fridrich, and T. Denemark, "Universal distortion function for steganography in an arbitrary domain," *EURASIP Journal on Information Security*, vol. 2014, no. 1, p. 1, 2014.



(a) HILL map (b) UM-HILL map (c) bias-HILL map

Fig. 4. Steganalytic performance of ES(mixture group), ES(HILL group) and ES(HILL and UNIWARD) under maxSRMd2 detection, each group is tested with the modification probability maps generated by its disturbing functions. (a)-(c) are the experimental results of each group.



(a) UNIWARD map (b) UM-UNIWARD map (c) bias-UNIWARD map

Fig. 5. Steganalytic performance of ES(mixture group), ES(UNIWARD group) and ES(HILL and UNIWARD) under maxSRMd2 detection, each group is tested with the modification probability maps generated by its disturbing functions.

[9] B. Li, M. Wang, J. Huang, and X. Li, "A new cost function for spatial image steganography," in *Image Processing (ICIP), 2014 IEEE International Conference on*. IEEE, 2014, pp. 4206–4210.

[10] W. Tang, H. Li, W. Luo, and J. Huang, "Adaptive steganalysis against wow embedding algorithm," in *Proceedings of the 2nd ACM workshop on Information hiding and multimedia security*. ACM, 2014, pp. 91–96.

[11] T. Denemark, V. Sedighi, V. Holub, R. Cogranne, and J. Fridrich, "Selection-channel-aware rich model for steganalysis of digital images," in *Information Forensics and Security (WIFS), 2014 IEEE International Workshop on*. IEEE, 2014, pp. 48–53.

[12] W. Tang, H. Li, W. Luo, and J. Huang, "Adaptive steganalysis based on embedding probabilities of pixels," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 4, pp. 734–745, 2016.

[13] T. Denemark, J. Fridrich, and P. Comesaña-Alfaro, "Improving selection-channel-aware steganalysis features," *Electronic Imaging*, vol. 2016, no. 8, pp. 1–8, 2016.

[14] T. Filler and J. Fridrich, "Gibbs construction in steganography," *IEEE Transactions on Information Forensics and Security*, vol. 5, no. 4, pp. 705–720, 2010.

[15] P. Bas, T. Filler, and T. Pevnỳ, "Break our steganographic system: The ins and outs of organizing BOSS," in *International Workshop on Information Hiding*. Springer, 2011, pp. 59–70.

[16] K. Chen, W. Zhang, H. Zhou, N. Yu, and G. Feng, "Defining cost functions for adaptive steganography at the microscale," in *Information Forensics and Security (WIFS), 2016 IEEE International Workshop on*. IEEE, 2016, pp. 1–6.

[17] J. Li, X. Yang, X. Liao, F. Pan, and M. Zhang, "A game-theoretic method for designing distortion function in spatial steganography," *Multimedia Tools and Applications*, vol. 76, no. 10, pp. 12 417–12 431, 2017.