Copyright © 2014 KSII

Recognizing F5-like stego images from multi-class JPEG stego images

Jicang Lu, Fenlin Liu and Xiangyang Luo
[1] Zhengzhou Information Science and Technology Institute
[2] State key Laboratory of Mathematical Engineering and Advanced Computing
Zhengzhou, Henan 450001 - China
[3] State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, Beijing 100093 - China
[e-mail: lujicang@sina.com]
*Corresponding author: Jicang Lu

Received June 6, 2014; revised July 31, 2014; revised September 15, 2014; accepted October 2, 2014; published November 30, 2014

Abstract

To recognize F5-like (such as F5 and nsF5) steganographic algorithm from multi-class stego images, a recognition algorithm based on the identifiable statistical feature (IDSF) of F5-like steganography is proposed in this paper. First, this paper analyzes the special modification ways of F5-like steganography to image data, as well as the special changes of statistical properties of image data caused by the modifications. And then, by constructing appropriate feature extraction sources, the IDSF of F5-like steganography distinguished from others is extracted. Lastly, based on the extracted IDSFs and combined with the training of SVM (Support Vector Machine) classifier, a recognition algorithm is presented to recognize F5-like stego images from images set consisting of a large number of multi-class stego images. A series of experimental results based on the detection of five types of typical JPEG steganography (namely F5, nsF5, JSteg, Steghide and Outguess) indicate that, the proposed algorithm can distinguish F5-like stego images reliably from multi-class stego images generated by the steganography mentioned above. Furthermore, even if the types of some detected stego images are unknown, the proposed algorithm can still recognize F5-like stego images correctly with high accuracy.

Keywords: Steganalysis, Recognition of stego images, F5, nsF5, Identifiable statistical feature

This research was supported by a research grant from the National Natural Science Foundation of China (Grant Nos. 61272489, 61379151 and 61302159), the Excellent Youth Foundation of Henan Province of China (Grant No. 144100510001), the Fundation of Science and Technology on Information Assurance Labaratory (Grant No. KJ-14-108), the Scientific and Technological Innovation Leading Talent of Zhengzhou (Grant No. 10LJRC182), and the Doctoral Dissertation Innovation Fund of Zhengzhou Information Science and Technology Institute (Grant No. BSLWCX201203).

http://dx.doi.org/10.3837/tiis.2014.11.028

1. Introduction

The techniques of steganography and steganalysis are now two important research directions in information security [1]. Steganography is a covert communication technique that embeds confidential messages into the redundant parts of multimedia files such as digital images and videos, and then transfers them through public communication channels [2]. Contrarily, steganalysis aims to discover and prevent such covert communication behaviors. The main research aspects include: existence detection of the confidential messages [3,4,5], estimation of the embedded message length [6,7,8], recognition of the types of stego images [9,10], determination of the stego key [11], etc. Among all the above aspects, the reliable recognition and determination of the types of steganography used in the detected stego imagesare the significant prerequisite to the other three aspects, and are also key issues for further steganalysis and forensics [12]. This paper focuses on the problem of recognizing F5-like (such as F5 and nsF5) [13,14] stego images from multi-class stego images (Referred to as MultiClsStegImgs) generated by multiple types of steganographic algorithms, which can also be considered as recognizing the F5-like steganography.

Existing researches on recognition of the types of stego images mainly focus on the classification of MultiClsStegImgs. These methods are usually designed under the condition that the steganographic algorithms set possibly used in stego images are completely known, and then classify different types of stego images based on binary classification and using only one type of blind steganalytic feature. The issue above was first researched in [15], which classified JPEG MultiClsStegImgs based on coefficient histograms and co-occurrence matrix features. The algorithm was improved in [9,16] by merging the features of extended histogram and Markov transition probability matrix. In [10], a multi-class classification algorithm was presented using the feature of probability density function (PDF) moments extracted from wavelet coefficients. In [17], a classification algorithm for MultiClsStegImgs generated by spatial domain steganography was proposed based on run-length histogram features extracted from the pixel difference matrix. Almost all the features used by the methods above are conventional blind steganalytic features, and other types of blind steganalytic features (such as the features used in [3,4,5]) may also be utilized to classify MultiClsStegImgs. However, from the brief overview of MultiClsStegImgs classification above, it is clearly seen that there are mainly three shortcomings as follows: 1) The features extracted by existing methods are mostly for blind steganalysis, and often unable to reflect the special modifications caused by a specific steganography to statistical properties of image data; 2) Existing methods must know the steganographic algorithms possibly used in stego images, and could not be utilized to the condition with unknown types of stego images; 3) The classifiers used by existing methods are generally trained based on all the types of stego images, so, when adding a brand new type of stego images, new classifiers must be trained.

For the detection of F5-like steganography, existing researches mainly concentrate on two aspects: existence detection of the secret messages and estimation of the embedded messages. The former mainly distinguish F5-like stego images and cover images based on statistical feature extraction and classifier training. The popular features include: coefficients histogram [15], co-occurrence matrix [9,15], Markov transition probability matrix [9,15], rich models (high-dimensional features) [4,5], etc. On the other detection aspect, the typical methods are: the algorithm in [6] based on cover image estimation and the least square analysis, the algorithm in [18] based on blind steganalytic features and the regression analysis, etc. In the

previous researches, we proposed an estimation algorithm based on relative entropy of the histogram for F5-like steganography [19], and an algorithm based on blind steganalytic features and ensemble learning [20]. However, these algorithms above only dedicate to determine the existence of secret messages or estimating the length of embedded messages, and there are few published researches on the recognition of F5-like stego images from MultiClsStegImgs.

For the recognition of F5-like stego images from multi-class stego images, the main contribution of this paper is: based on the analysis and extraction of the identifiable statistical feature (Referred to as IDSF) for F5-like steganography, a recognition algorithm is proposed to distinguish F5-like stego images from MultiClsStegImgs. Firstly, the special modification ways of F5-like steganography to image data during embedding messages are analyzed. Then, the modifications are depicted based on the changes of differences between neighboring DCT (Discrete Cosine Transform) coefficients as well as neighboring pixels, and the IDSF of F5-like steganography distinguished from others is extracted. At last, the recognition algorithm for F5-like stego images is presented based on IDSF. The efficiency of the proposed algorithm is verified by a series of experiments, which indicate that the performance of the proposed algorithm is superior to that of existing typical classification algorithm.

The rest of this paper is organized as follows: Section 2 will analyze the changing rules of F5-like steganography to the cover. In Section 3, according to the special modification ways analyzed in Section 2, the IDSF of F5-like steganography is analyzed and extracted to distinguish them from other types of steganography, and the recognition algorithm for F5-like stego images is also presented. Section 4 will report the experiments verifying the efficiency of the proposed recognition algorithm. The paper is concluded in Section 5.

2. Modification rules of F5-like steganography to image

F5-like steganography is a class of popular steganographic algorithms, and currently more secure algorithms. The principal characteristics of these algorithms are:

(i) Modification ways during embedding. When the coefficients have to be modified, it is not to deal with the LSB (least significant bit) directly, but rather to subtract 1 from the absolute coefficient while preserving the sign to embed the message bits.

(ii) Determination of the coefficients to be modified. To improve the security, the matrix encoding is implemented to improve the embedding efficiency (to embed more messages with as fewer changes as possible), and the permutation straddling mechanism is employed to make the changes be randomly distributed to the whole image.

At present, the typical steganographic algorithms designed based on the strategies above include F5 and nsF5. Denote C as the DCT coefficient matrix of an image, and $C_{k,l,u,v}$ as the coefficient in position (u,v) of the block in the k-th row and l-th column, after embedding, it is $C'_{k,l,u,v}$. Then, the modification ways of the coefficients during embedding by F5-like steganography can be depicted as following expressions:

$$\begin{cases} C'_{k,l,u,v} = C_{k,l,u,v} & LSB(C_{k,l,u,v}) = b \\ C'_{k,l,u,v} = sign(C_{k,l,u,v}) * (|C_{k,l,u,v}| - 1) & LSB(C_{k,l,u,v}) \neq b \end{cases} \quad (1)$$

where, b is the embedded message bit; $|\bullet|$ is used to calculate absolute value of the data; $LSB(C_{k,l,u,v})$ is a function to calculate the LSB of $C_{k,l,u,v}$. Denote h_c and h'_c as the

histograms of coefficients with value c before and after embedding, respectively, then, when the average changing rate of the coefficients is β, the relationship of the histograms before and after embedding is:

$$\begin{cases} h'_c = h_c + \beta h_{c+1} + \beta h_{c-1} & c = 0 \\ h'_c = h_c - \beta h_c + \beta h_{c+1} & c > 0 \\ h'_c = h_c - \beta h_c + \beta h_{c-1} & c < 0 \end{cases} \tag{2}$$

For F5 steganography, the coefficients with value 0 will not be used to embed messages. When the message bit is hidden into the coefficient with value 1 or -1 and the coefficient is changed to 0, this embedding is invalid, and the bit will be re-hidden into the next nonzero coefficients. For natural JPEG images, the coefficients histograms satisfy characteristic: $\cdots < h_{-3} < h_{-2} < h_{-1} < h_0 > h_1 > h_2 > h_3 > \cdots$, then, in Eq.(2), there is $h'_0 > h_0$, and $h'_c < h_c$ for $c \neq 0$. Therefore, F5 will lead to the phenomenon of serious histogram shrinkage.

Jessica *et al.* improved the F5 steganography using wet paper code [14], and called the improved algorithm as nsF5. When the coefficients with value 1 or -1 are changed to 0, nsF5 will not re-hide the secret message bit, which avoids invalid embedding and abate the phenomenon of histogram shrinkage. In **Fig. 1**, the coefficients histograms of the cover image (Cover), the F5 and nsF5 stego images are plotted for comparison.

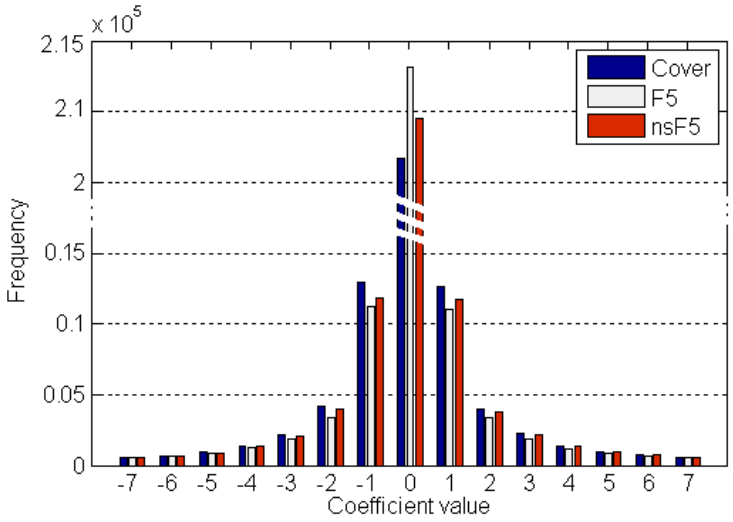

Fig. 1. The comparison of DCT coefficients histograms of the Cover image, F5 and nsF5 stego images.

It can be seen from **Fig. 1** and Eq.(2) that, although nsF5 steganography avoids invalid embedding, but does not solve the problem of histogram shrinkage. The reader might refer to [13,14,19] for a more detailed introduction of F5-like steganography. In what follows, based on the phenomenon of histogram shrinkage and embedding changing rules of F5-like steganography, the sensitive statistical features will be analyzed and extracted to capture the special phenomenon and modification ways.

3. IDSF based recognition of F5-like stego images

Generally, different steganography will produce different modifications to image data, as well as different influences to statistical properties of different feature extraction sources. Therefore, different statistical features should be extracted for corresponding steganography to achieve more effective steganalysis. For instance, for two types of typical steganalytic features characteristic function (CF) moments and probability density function (PDF) moments extracted from wavelet coefficients, their performances to different steganography were theoretically analyzed and compared in [21,22]. For the reliable recognition of a specific type or class of stego images from multi-class stego images, it is significant to extract statistical features that be sensitive to the specific type or class of steganography and not (or less) sensitive to others. To extract features with above statistical properties, one should firstly construct appropriate feature extraction sources according to the special message embedding and modification ways of the specific steganography different from others, and then extract sensitive features that could capture the special modification ways. In this paper, this type of feature is called IDSF of the steganography distinguished from others.

3.1 Extracting IDSF

According to the embedding and modification ways of F5-like steganography introduced in Section 2, this paper will analyze the influencing of the embedding changes to statistical properties of data in DCT domain and spatial domain. And then, by constructing appropriate feature extraction sources, extract sensitive features that could capture the special modification ways of F5-like steganography. At last, take the sensitive features as the IDSF of F5-like steganography distinguished from others.

3.1.1 Extracting IDSF from DCT domain

It can be seen from the introduction of F5-like steganography in Section 2 that, if the DCT coefficient changes after embedding, absolute value of the coefficient will decrease by 1, i.e., the positive coefficients will decrease by 1, and the negative coefficients will increase by 1. Therefore, for the difference of neighboring coefficients under certain conditions, the sign and changing trend after embedding can also be determined. Let

$$\Delta C = C_{k,l,u,v} - C_{k+\Delta_1,l+\Delta_2,u+\Delta_3,v+\Delta_4} \tag{3}$$

$$\Delta C' = C'_{k,l,u,v} - C'_{k+\Delta_1,l+\Delta_2,u+\Delta_3,v+\Delta_4} \tag{4}$$

$$\nabla C = \Delta C - \Delta C' \tag{5}$$

where, $(\Delta_1,\Delta_2,\Delta_3,\Delta_4) \in \{(0,0,1,0),(0,0,0,1),(0,0,1,1),(1,0,0,0),(0,1,0,0),(1,1,0,0)\}$ are used to control the relative locations and directions (horizontal, vertical or diagonal) of the neighboring coefficients. For example, when $(\Delta_1,\Delta_2,\Delta_3,\Delta_4)=(0,0,1,0)$, ΔC denotes the difference between neighboring coefficients along the vertical direction of intra-block, and when $(\Delta_1,\Delta_2,\Delta_3,\Delta_4)=(0,1,0,0)$, ΔC denotes the difference between neighboring coefficients along the horizontal direction of inter-block. Let

$$C_{Multi} = C_{k,l,u,v} * C_{k+\Delta_1,l+\Delta_2,u+\Delta_3,v+\Delta_4} \tag{6}$$

$$C_{Plus} = C_{k,l,u,v} + C_{k+\Delta_1,l+\Delta_2,u+\Delta_3,v+\Delta_4} \tag{7}$$

Then, according to the principle of F5-like steganography, the changing trend of variable ∇C in Eq.(5) after embedding changes are as follows:

$$\begin{cases} case\,1: & |\nabla C| < 0 \ or \ |\nabla C| > 0 & C_{Multi} > 0 \\ case\,2: & |\nabla C| < 0 & C_{Multi} < 0 \\ case\,3: & |\nabla C| < 0 & C_{Multi} = 0 \ and \ C_{Plus} \neq 0 \end{cases} \tag{8}$$

It can be seen from Eq.(8) that, for the case 1 ($C_{Multi} > 0$), the changing trend of the coefficients difference after embedding could not be determined according to the coefficient values only; while for case 2 ($C_{Multi} < 0$) and case 3 ($C_{Multi} = 0$ *and* $C_{Plus} \neq 0$), absolute value of the coefficients difference will decrease after embedding change. Whereas for other types of steganography such as JSteg[1], Outguess[2] and Steghide[3], no matter what the relationship between signs of the neighboring coefficients is, the change of the difference between neighboring coefficients could not be determined at all according to the value of the coefficients only. The phenomenon above is exactly the special characteristic of F5-like steganography different from other types of steganography mentioned above.

Therefore, if one wants to extract the IDSF of F5-like steganography distinguished from others, it is important to construct feature extraction sources according to the coefficients under case 2 and case 3 in Eq.(8), and then the sensitive features can be extracted. In this paper, the histogram and co-occurrence matrix will be calculated from the coefficients differences under case 2 and case 3 in Eq.(8). For natural images, the coefficients histogram is symmetry about 0, and the histogram of coefficients differences is also symmetry about 0. Then, the features will be calculated from absolute values of the coefficients differences.

First, calculate the histogram feature. Denote D_{inter}^{N0} and D_{intra}^{N0} as the set of differences between two neighboring coefficients with opposite signs (i.e., $C_{Multi} < 0$) and with only one nonzero (i.e. $C_{Multi} = 0$ and $C_{Plus} \neq 0$) in the inter- and intra-block, respectively, which contains neighboring coefficients along the horizontal, vertical and diagonal directions in all. The histograms H_{inter}^{N0} and H_{intra}^{N0} of the differences sets above are calculated as follows:

$$H_{inter}^{N0}(d) = \frac{\sum_i \delta\left(\left|D_{inter}^{N0}(i)\right|, d\right)}{\sum_d H_{inter}^{N0}} \tag{9}$$

$$H_{intra}^{N0}(d) = \frac{\sum_i \delta\left(\left|D_{intra}^{N0}(i)\right|, d\right)}{\sum_d H_{intra}^{N0}} \tag{10}$$

where, $1 \le d \le T$, and T is the threshold for differences histogram. $\delta(\Lambda_1, \Lambda_2) = 1$ if and only if $\Lambda_1 = \Lambda_2$; otherwise $\delta(\Lambda_1, \Lambda_2) = 0$.

Then, calculate the co-occurrence matrix features. In order to depict the special changes of

[1] JSteg: Available at http://zooid.org/~paul/crypto/jsteg. 2014
[2] Outguess: Available at http://www.outguess.org. 2014
[3] Steghide: Available at http://steghide.sourceforge.net. 2014

F5-like steganography different from others better, if two neighboring differences include the difference in the case 1 in Eq.(8), then, this pair of differences will not be included for calculation of the co-occurrence matrix. Let

$$\Delta C_2 = C_{k+\Delta_1,l+\Delta_2,u+\Delta_3,v+\Delta_4} - C_{k+2\Delta_1,l+2\Delta_2,u+2\Delta_3,v+2\Delta_4} \tag{11}$$

$$C_{Multi,2} = C_{k+\Delta_1,l+\Delta_2,u+\Delta_3,v+\Delta_4} * C_{k+2\Delta_1,l+2\Delta_2,u+2\Delta_3,v+2\Delta_4} \tag{12}$$

where $(\Delta_1,\Delta_2,\Delta_3,\Delta_4)$ are the same to that in Eqs.(3)~(5). Based on the expressions in Eqs.(3), (6) and (12), calculate the co-occurrence matrix for differences of neighboring coefficients in inter- and intra-block using the following equation.

$$F^{CM}(d_1,d_2) = \frac{\sum_{k,l,u,v,} \delta(|\Delta C|,d_1)\delta(|\Delta C_2|,d_2)}{\sum_{d_1,d_2} F^{C}(d_1,d_2)}, \quad C_{Multi} \le 0 \text{ and } C_{Multi,2} \le 0 \tag{13}$$

where, $0 \le d_1,d_2 \le T$; when $(\Delta_1,\Delta_2,\Delta_3,\Delta_4) \in \{(1,0,0,0),(0,1,0,0),(1,1,0,0)\}$, the results calculated by Eq.(13) is the co-occurrence matrix for differences of neighboring coefficients in inter-blocks, and is denoted as F^{CM}_{inter}; when $(\Delta_1,\Delta_2,\Delta_3,\Delta_4) \in \{(0,0,1,0),(0,0,0,1),(0,0,1,1)\}$, the result calculated by Eq.(13) is the co-occurrence matrix for differences of neighboring coefficients in intra-block, and is denoted as F^{CM}_{intra}.

In addition, referring to the method of image calibration in [15,16], the corresponding calibrated features of the detected images will also be extracted in this paper. Firstly, crop the upmost 4 rows and leftmost 4 columns. And then, quantize and compress the cropped image with the original quantization matrix, the corresponding calibrated image (i.e. reference image of the original image) can be obtained. At last, extract the histograms and co-occurrence matrixes of differences from the calibrated image, and denote them as $H^{cal,N0}_{inter}$, $H^{cal,N0}_{intra}$, $F^{cal,CM}_{inter}$ and $F^{cal,CM}_{intra}$, respectively.

In summary, the DCT part of IDSF $F^{DCT}_{F5-like}$ extracted based on various pair relationships of neighboring DCT coefficients for F5-like steganography is as follows:

$$F^{DCT} = H^{N0}_{inter} \cup H^{N0}_{intra} \cup H^{cal,N0}_{inter} \cup H^{cal,N0}_{intra} \cup F^{CM}_{inter} \cup F^{CM}_{intra} \cup F^{cal,CM}_{inter} \cup F^{cal,CM}_{intra} \tag{14}$$

3.1.2 Extracting IDSF from spatial domain

It can be seen from the analysis in Section 2 that, the coefficients modified by F5-like steganography will change towards 0. Then, based on this characteristic and the mutual expression between DCT coefficients and pixels, whether or not the deterministic changing of the statistical properties of pixels as well as neighboring pixels? The answer is positive.

Generally, JPEG images are compressed and saved by blocks with size 8×8. For each block, the IDCT (Inverse DCT) expression to transform DCT coefficients into pixels is:

$$f(i,j) = \frac{1}{4}\sum_{u=0}^{7}\sum_{v=0}^{7} c(u)c(v)F(u,v)\cos\frac{(2i+1)u\pi}{16}\cos\frac{(2i+1)v\pi}{16} \tag{15}$$

where, $f(i,j)$ denotes the pixel in position (i,j) of the block in the spatial domain ($i,j \in \{0,1,...,7\}$). $F(u,v)$ denotes the DCT coefficient before quantization, and is in the position (u,v) of the block in the DCT domain. $c(u)$ and $c(v)$ are as follows:

$$\begin{cases} c(u) = \begin{cases} 1/\sqrt{2} & u = 0 \\ 1 & u \neq 0 \end{cases} \\ c(u) = \begin{cases} 1/\sqrt{2} & v = 0 \\ 1 & v \neq 0 \end{cases} \end{cases} \tag{16}$$

For neighboring pixels of the local area inside the image block, their differences and corresponding changes after F5-like steganography are:

$$\Delta f_{0,0}^{0,1} = f(0,1) - f(0,0) = \frac{1}{4}\sum_{u=0}^{7}\sum_{v=0}^{7} c(u)c(v)F(u,v)\cos\frac{u\pi}{16}\left(-2\sin\frac{v\pi}{8}\sin\frac{v\pi}{16}\right) \tag{17}$$

$$\Delta f_{0,0}^{1,0} = f(1,0) - f(0,0) = \frac{1}{4}\sum_{u=0}^{7}\sum_{v=0}^{7} c(u)c(v)F(u,v)\cos\frac{v\pi}{16}\left(-2\sin\frac{u\pi}{8}\sin\frac{u\pi}{16}\right) \tag{18}$$

In Eqs.(17) and (18), if $F(u,v) > 0$, then $\Delta f_{0,0}^{0,1} < 0$ and $\Delta f_{0,0}^{1,0} < 0$. After embedding change, $F(u,v)$ will decrease, $\Delta f_{0,0}^{0,1}$ and $\Delta f_{0,0}^{1,0}$ will both increase while the absolute values of them will both decrease. Contrarily, if $F(u,v) < 0$, then $\Delta f_{0,0}^{0,1} > 0$ and $\Delta f_{0,0}^{1,0} > 0$. After embedding change, $F(u,v)$ will increase, $\Delta f_{0,0}^{0,1}$ and $\Delta f_{0,0}^{1,0}$ will both decrease while the absolute values of them will also decrease.

It can be seen from above that, the histograms of differences of neighboring pixels in Eqs.(17)~(18) will shrink after embedding change for F5-like steganography. However, for other types of JPEG steganography such as JSteg, Outguess and Steghide, the shrinkage above will not happen, and the changes of neighboring pixels differences in Eqs.(17)~(18) could not be determined according to signs of the coefficients only, either. The changes above are exactly the special characteristic of F5-like steganography different from others. Therefore, based on the pixels differences calculated in Eqs.(17) and (18), the difference histogram will be calculated in this paper and taken as IDSF of F5-like steganography.

Denote D_{Pix}^{P00} as the set of differences $\Delta f_{0,0}^{0,1}$ and $\Delta f_{0,0}^{1,0}$ corresponding to positive coefficients in all the blocks, and D_{Pix}^{N00} as the set of differences $\Delta f_{0,0}^{0,1}$ and $\Delta f_{0,0}^{1,0}$ corresponding to all the negative coefficients. Then, calculate the histogram feature of the differences above as follows:

$$H_{Pdiff}^{00}(d) = \frac{\sum_i \left(\delta\left(D_{Pix}^{P00}(i),d\right) + \delta\left(D_{Pix}^{N00}(i),d\right)\right)}{\sum_d H_{Pdiff}^{00}} \tag{19}$$

where, $-T \le d \le T$, and T is the threshold for pixel differences. For the reference image by cropping the upmost 4 rows and leftmost 4 columns, the two histogram features are also calculated and denoted as $H_{Pdiff}^{cal,00}$. At last, the IDSF of F5-like steganography extracted from spatial domain can be expressed as follows:

$$F^{Pix} = H_{Pdiff}^{00} \bigcup H_{Pdiff}^{cal,00} \tag{20}$$

In summary, based on the analysis of modifications and influences caused by F5-like steganography to the data in DCT domain and spatial domain, the IDSF of F5-like steganography distinguished from others can be obtained as follows:

$$F = F^{DCT} \bigcup F^{Pix} \tag{21}$$

3.2 Recognition algorithm

In practical steganalysis, the detector usually could not know exactly the steganographic algorithms possibly used in the MultiClsStegImgs. Maybe only one or some of the steganographic algorithms used in parts of the stego images could be determined by the detector, but the detector cannot determine exactly the algorithm used in any stego image. For this special condition, the recognition algorithm for F5-like steganography will be presented in this paper, and the main steps are as follows:

Step 1: Extracting features from the detected images. According to Eqs.(9), (10), (13) and (19) as well as the feature extraction threshold *T*, extract features from DCT coefficients and spatial pixels, respectively.

Step 2: Constructing the reference image. Using the image calibration method based on cropping of 4 rows and columns, construct the reference image of the detected image.

Step 3: Extracting features from the reference image. According to Eqs.(9), (10), (13) and (19) as well as the feature extraction threshold *T*, extract features from DCT coefficients and spatial pixels of the reference image, respectively.

Step 4: Generating IDSF of F5-like steganography. According to Eqs.(14) and (20), process the features extracted in Step 1 and Step 3, and then generate the IDSF of F5-like steganography.

Step 5: Constructing images set and training classifier. First, construct images set for classifier training based on the partially known steganographic algorithms and F5-like steganography. Then, extract features of the constructed images according to Steps 1~4, and select an appropriate classifier training algorithm to train the classification and recognition model for F5-like steganography.

Step 6: Classifying and recognizing the F5-like stego images. Using the model trained in Step 5, classify the features extracted in Step 4, the detected images can be classified as F5-like stego images or non-F5-like stego images (images generated by other types of steganography).

4. Experimental results and analysis

4.1 Experiments setup

In this section, the proposed algorithm will be experimentally verified and analyzed based on a well-known images database BossBase-1.01[4] in current steganalysis. The original 5000 images were taken from this images database in grayscale PGM format with size 512×512. Firstly, convert all the PGM images into JPEG images saved with a quality factor 75; Then, construct stego images using 5 types of typical JPEG steganography F5, nsF5, JSteg, Steghide and Outguess with four payloads (message bits carried by per nonzero AC DCT coefficients), namely 0.1, 0.2, 0.3 and 0.5 bpnzAC (bits per nonzero AC DCT coefficient). In total, the images set contained 5000×4×5 = 100000 stego images (which consists of the MultiClsStegImgs set) and 5000 cover images, and the details were listed in **Table 1**.

Table 1. Illustration of experimental images set

Image type	Payload (bpnzAC)	Total number
Cover images	0	5000
F5 stego images	0.1, 0.2, 0.3, 0.5	5000×4 = 20000
nsF5 stego images	0.1, 0.2, 0.3, 0.5	5000×4 = 20000
JSteg stego images	0.1, 0.2, 0.3, 0.5	5000×4 = 20000
Steghide stego images	0.1, 0.2, 0.3, 0.5	5000×4 = 20000
Outguess stego images	0.1, 0.2, 0.3, 0.5	5000×4 = 20000

According to the IDSF extraction methods proposed in Section 3 for F5-like steganography, extract features of the experimental stego images set. For feature extraction, it can be seen from Eqs.(13) and (19) that, dimensionality of the feature will increase noticeably as the threshold T increases. The high dimensionality of the feature usually leads to the problem of "curse of dimensionality" and reduces the efficiency of the detection algorithm. It is very important to control the dimensionality of the extracted features. For natural JPEG images, many researches have shown that, an overwhelming majority of the differences between neighboring coefficients and pixels are distributed in a narrow range nearby zero, and the steganalytic features are usually extracted by setting the threshold T as a small value such as 5 [9,15]. If the threshold is larger than 5, then, the features extracted from differences larger than 5 may be redundant ones, and may not play an active role in the classification. If the threshold is smaller than 5, some of the changes might not be captured. Then, the extracted features may not achieve the best classification results. Therefore, the threshold T is set to 5 in this paper, then, the dimensionality of the feature extracted based on Eq.(14) from DCT coefficients is 164, and the dimensionality of the feature extracted based on Eq.(20) from spatial pixels is 22. The total dimensionality of IDSF proposed in this paper for F5-like steganography is 186.

On the application of the classifier training method, current steganalytic algorithms [9,17] based on low-dimensional statistical features usually use the support vector machine (SVM) [23] classifier training method with Gaussian kernel. This classifier training method projects the low-dimensional features into high-dimensional feature space, and then finds a proper

[4] BOSS: Available: http://exile.felk.cvut.cz/boss/BOSSFinal/. 2014.

decision boundary to classify different samples, which usually achieves better detection performance. Therefore, in the experiments of this paper, the SVM classifier training method with a Gaussian kernel was used in the proposed recognition algorithm. The experiments were repeated ten times, and the average values of the ten results were taken as the final detection results for comparison. In addition, for the reason that the detectors may not know clearly the steganographic algorithms possibly used in MultiClsStegImgs in practical steganalysis, therefore, this paper will analyze and test the performance of the proposed IDSF based recognition algorithm in the condition that only one type of steganography is known to the detector, which is used to recognize F5-like stego images.

4.2 Testing on recognition of F5-like stego images

As introduced in Section 1, existing recognition algorithms for stego images were usually designed based on the blind steganalytic features with better performance. The algorithm in [9] is a typical classification algorithm for multi-class stego images. The features used by [9] are 548 dimensional CCPEV features, which are extracted based on image calibration [16] and have better performance. The SVM with Gaussian kernel is also used by [9]. However, when only one type or class of stego images is needed to be recognized, and types of some detected images are unknown, the original classification framework in [9] cannot be applied any more. The multi-class classifiers could not be trained using the feature CCPEV in [9] either. In the following, the experiments will be carried out in this condition.

When some of the steganographic algorithms possibly used in the multi-class stego images set are unknown, based on the proposed IDSF of F5-like steganography and the feature in [9], the classification and recognition procedure presented in Subsection 3.2 will be applied to verify the efficiency of the proposed recognition algorithm as well as the IDSF. The detection based on CCPEV and the proposed recognition procedure is referred to as "CCPEV with the proposed procedure". And the detection based on the proposed IDSF and the proposed procedure is referred to as "F5likeIdf (Proposed)". Now, the classifier will be trained based on F5-like stego images and the known type of stego images only. Other types of stego images will not be included in training classifier. And then, classify the multi-class stego images as F5-like stego images and non-F5-like stego images. For each testing, 3000 stego images of each algorithm are randomly selected for training, and the remaining 2000 ones are used for testing. Suppose that only JSteg is known to the detector, the classifier used for testing will be trained based on F5 stego images and nsF5 stego images versus JSteg stego images, respectively. The corresponding 2000 stego images in other types (Steghide and Outguess) of stego images are also used for testing. Then, the numbers of stego images under four payloads (0.1, 0.2, 0.3 and 0.5 bpnzAC) involved in the training and testing are $3000\times4\times2 = 24000$ and $2000\times4\times5 = 40000$, respectively. The detection results of the proposed algorithm (i.e. F5likeIdf (Proposed)) and the algorithm with the feature CCPEV (i.e. CCPEV with the proposed procedure) are comparatively listed in **Table 2** and **Table 3**, respectively. At the same time, the ROC (Receiver Operation Characteristic) curves corresponding to **Table 2** and **Table 3** are plotted in **Fig. 2.** The ROC curve is used to depict the correct detection probabilities of a detection algorithm corresponding to various false alarm probabilities. It can reflect the overall performance of a steganalytic algorithm, and is a popular method for performance comparison in current steganalysis. In **Fig. 2,** the false alarm probability denotes the percentage that other types of stego images are misclassified as F5-like stego images, and the detection probability denotes the percentage that the F5-like stego images are correctly classified corresponding to a specific false alarm probability. For example, for a given false alarm probability in the horizontal axis, the corresponding value in the vertical axis denotes

the detection probability of the steganalytic algorithm. The larger the detection probability is, the more superior the steganalytic algorithm will be.

It can be seen from the results in **Table 2** and **Table 3** that, when there is only Outguess is known, utilizing the procedure in Subsection 3.2 for recognition of F5-like stego images, the recognition results based on the proposed IDSF are superior to that based on the CCPEV feature in [9]. Especially, when the embedding ratio is 0.1 bpnzAC, the algorithm based on the feature CCPEV in [9] misclassified almost all the stego images generated by Steghide and Outguess as F5-like stego images. At the same time, it can be seen from **Fig. 2** that, for each false alarm probability in the horizontal axis, the detection probabilities of the proposed IDSF F5likeIdf are all higher than that of the typical feature CCPEV, which indicate that, the overall performance of the algorithm based on the proposed IDSF outperforms the algorithm based on the feature CCPEV. The results above indicate that, using the IDSF based recognition algorithm proposed in this paper, only one classifier is enough to achieve reliable recognition, which can verify the efficiency of the proposed IDSF and recognition algorithm for recognition of F5-like steganography.

Table 2. The probability of each type of stego images classified as F5-like stego images based on the classifier trained using F5 and JSteg stego images. For F5-like (F5 and nsF5) stego images, the results are correct detection accuracy, while for non-F5-like (JSteg, Steghide and Outguess) stego images, the results are error detection accuracy.

payload (bpnzAC)	Detection algorithm	F5	nsF5	non-F5-like = (JSteg + Steghide + Outguess) / 3
0.1	CCPEV with the proposed scheme	99.93	99.91	60.46 = (0.03 + 87.92 + 93.42) / 3
	F5likeIdf (Proposed)	99.85	99.67	11.86 = (0.23 + 7.42 + 27.93) / 3
0.2	CCPEV with the proposed scheme	99.95	99.93	14.72 = (0 + 21.37 + 22.78) / 3
	F5likeIdf (Proposed)	99.87	99.67	0.42 = (0 + 0.15 + 1.12) / 3
0.3	CCPEV with the proposed scheme	99.95	99.92	0.63 = (0 + 0.70 + 1.18) / 3
	F5likeIdf (Proposed)	99.97	99.70	0.17 = (0 + 0.05 + 0.45) / 3
0.5	CCPEV with the proposed scheme	99.98	99.88	0.22 = (0 + 0.08 + 0.58) / 3
	F5likeIdf (Proposed)	99.97	99.57	0.09 = (0 + 0.02 + 0.27) / 3

Table 3. Based on the classifier trained using nsF5 and Outguess stego images, the probability of each type of stego images classified as F5-like stego images. For F5-like (F5 and nsF5) stego images, the results are correct detection accuracy, while for non-F5-like (JSteg, Steghide and Outguess) stego images, the results are error detection accuracy.

payload (bpnzAC)	Detection algorithm	F5	nsF5	non-F5-like = (JSteg + Steghide + Outguess) / 3
0.1	CCPEV with the proposed scheme	99.97	99.90	59.33 = (0.12 + 86.80 + 91.07) / 3
	F5likeIdf (Proposed)	99.90	99.82	13.02 = (0.58 + 8.38 + 30.10) / 3
0.2	CCPEV with the proposed scheme	99.97	99.93	6.62 = (0 + 10.08 + 9.78) / 3
	F5likeIdf (Proposed)	99.92	99.82	0.49 = (0 + 0.25 + 1.23) / 3
0.3	CCPEV with the proposed scheme	99.97	99.93	0.15 = (0 + 0.22 + 0.23) / 3
	F5likeIdf (Proposed)	99.97	99.85	0.19 = (0 + 0.08 + 0.50) / 3
0.5	CCPEV with the proposed scheme	99.98	99.93	0.08 = (0 + 0.03 + 0.20) / 3
	F5likeIdf (Proposed)	99.98	99.83	0.07 = (0 + 0 + 0.22) / 3

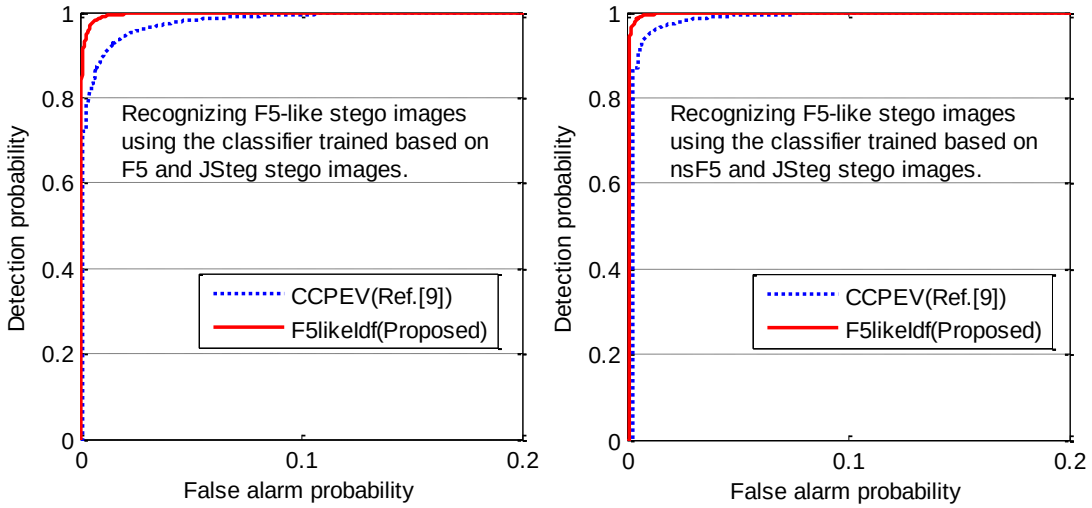

Fig. 2. When the classifier is trained based on F5 and JSteg or nsF5 and JSteg, the comparison of ROC curves of the two detection algorithms.

In addition, when only the algorithm Steghide or Outguess is known, the classification results of two detection algorithms are similar to the results listed in **Table 2** and **Table 3**, and the results will not be listed in detail. In the following, performances of the two detection algorithms will be compared using the average detection accuracy and time consuming (the

time consumed by classifier training and testing during detection) under every condition. The results calculated above are comparatively listed in **Table 4** and **Table 5**, respectively.

Table 4. When the classifier is trained based on F5 and one known steganographic algorithm, the comparison of average detection accuracy (%) and average time consuming (sec) of the two detection algorithms.

The known algorithm	Average detection accuracy		Average time consuming	
	CCPEV with the proposed scheme	F5likeIdf (Proposed)	CCPEV with the proposed scheme	F5likeIdf (Proposed)
JSteg	88.57	98.02	613.98	134.40
Steghide	99.74	99.53	704.96	150.28
Outguess	99.64	99.51	711.58	178.32

Table 5. When the classifier is trained based on nsF5 and one known steganographic algorithm, the comparison of average detection accuracy (%) and average time consuming (sec) of the two detection algorithms.

The known algorithm	Average detection accuracy		Average time consuming	
	CCPEV with the proposed scheme	F5likeIdf (Proposed)	CCPEV with the proposed scheme	F5likeIdf (Proposed)
JSteg	90.05	97.89	634.14	144.86
Steghide	99.67	99.47	719.80	165.60
Outguess	99.58	99.63	722.37	209.11

The results in **Table 4** and **Table 5** indicate further that, when only the steganographic algorithm JSteg is known, that is to say, the Steghide and Outguess stego images are not included in training classifier and are unknown to the classifier, the average detection accuracy and time consuming based on the proposed IDSF F5likeIdf for F5-like steganography are obviously superior to that based on the existing feature CCPEV. Similarly, when only Steghide or Outguess is known, although the average detection accuracies based on the two features are comparative with each other, however, it can be seen from the comparison of average time consuming that, the time consuming of the proposed IDSF F5likeIdf with a low dimensional feature is far shorter than that of CCPEV feature. The results analyzed above verify further that, when some types of the stego images are unknown, the proposed IDSF and recognition algorithm could still recognize F5-like stego images effectively.

5. Conclusion

For the reliable recognition of F5-like stego images from MultiClsStegImgs, this paper proposed a recognition algorithm based on IDSF of F5-like steganography. According to the special modification ways of F5-like steganography different from others, the IDSF of F5-like steganography was presented. Then, utilize the proposed IDSF and combine with SVM, corresponding recognition algorithms were presented. At last, a series of experimental results based on the classification of stego images generated by 5 types of typical JPEG

steganography indicated that: the F5-like stego images could be recognized reliably from MultiClsStegImgs based on the proposed IDSF. Although the types of some stego images were unknown, the IDSF based recognition algorithm could still recognize F5-like stego images reliably. Compared with existing typical detection algorithms, the proposed algorithms could achieve higher recognition accuracy as well as faster detection speed.
The IDSF could provide a new basis for recognition of a specific type or class of stego images. However, this paper only discussed the extraction of IDSF for F5-like steganography, but on the aspects of efficient classification of F5-like stego images and recognition of other various types of stego images, it is needed to be researched further.

References

[1] Fabien A. P. Petitcolas, Ross J. Anderson and Markus G. Kuhn, "Information hiding—A survey," in *Proc. of the IEEE, special issue on protection of multimedia content*, vol. 87, no. 7, pp. 1062-1078, July, 1999. Article (CrossRef Link).

[2] Abbas Cheddad, Joan Condell, Kevin Curran and Paul Mc Kevitt, "Digital image steganography: Survey and analysis of current methods," *Signal Processing*, vol. 90, no. 3, pp. 727-752, March, 2010. Article (CrossRef Link).

[3] Xiangyang Luo, Daoshun Wang, Ping Wang and Fenlin Liu, "A review on blind detection for image steganography," *Signal Processing*, vol. 88, no. 9, pp. 2138-2157, September, 2008. Article (CrossRef Link).

[4] Jan Kodovský and Jessica Fridrich, "Rich models for steganalysis of digital images," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 3, pp. 868-882, August, 2012. Article (CrossRef Link).

[5] Vojtěch Holub, Jessica Fridrich, "Random projections of residuals for digital image steganalysis," *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 12, pp. 1996-2006, December, 2013. Article (CrossRef Link).

[6] Jessica Fridrich, Miroslav Goljan and Dorin Hogea, "Steganalysis of JPEG Images: Breaking the F5 Algorithm," in *Proc. of 5th Int. Workshop on Information Hiding, Springer Lecture Notes in Computer Science*, vol. 2578, pp. 310-323, October 7-9, 2002. Article (CrossRef Link).

[7] Jessica Fridrich, Miroslav Goljan and Dorin Hogea, "Quantitative steganalysis of digital Images: Estimating the secret message length," *Multimedia Systems*, vol. 9, no. 3, pp. 288-302, November, 2003. Article (CrossRef Link).

[8] Chunfang Yang, Fenlin Liu, Shiguo Lian, Xiangyang Luo and Daoshun Wang, "Weighted stego-image steganalysis of messages hidden into each bit plane," *The Computer Journal*, vol. 55, no. 6, pp. 717-727, June, 2012. Article (CrossRef Link).

[9] Tomáš Pevný and Jessica Fridrich, "Multiclass detector of current steganographic methods for JPEG format," *IEEE Transactions on Information Forensics and Security*, vol. 3, no. 4, pp. 635-650, December, 2008. Article (CrossRef Link).

[10] Ping Wang, Fenlin Liu, Guodong Wang, Yifeng Sun and Daofu Gong, "Multi-class steganalysis for JPEG stego algorithms," in *Proc. of 15th IEEE Int. Conf. on Image Processing*, pp. 2076-2079, October 12-15, 2008. Article (CrossRef Link).

[11] Jessica Fridrich, Miroslav Goljan, David Sorkal and Taras Holotyak, "Forensic steganalysis: determining the stego key in spatial domain steganography," in *Proc. of SPIE, Security, Steganography, and Watermarking of Multimedia Contents VII*, SPIE 5681, pp. 631-642, January 17-21, 2005. Article (CrossRef Link).

[12] Andrew D. Ker, Patrick Bas, Rainer Böhme, Rémi Cogranne, Scott Craver, Tomáš Filler, Jessica Fridrich and Tomáš Pevný, "Moving steganography and steganalysis from the laboratory into the real world," in *Proc. of 1st ACM Workshop on Information Hiding and Multimedia Security*, pp. 45-58, Jun. 17-19 2013. Article (CrossRef Link).

[13] Andreas Westfeld, "F5—A steganographic algorithm—High capacity despite better steganalysis," in *Proc. of 4th Int. Workshop on Information Hiding*, *Springer Lecture Notes in Computer Science*,

vol. 2137, pp. 289-302, April 25-27, 2001. Article (CrossRef Link).
[14] Jessica Fridrich, Tomáš Pevný and Jan Kodovský, "Statistically undetectable JPEG steganography: Dead ends challenges, and opportunities," in *Proc. of 9th ACM Workshop on Multimedia and Security*, pp. 3-14, September 20-21, 2007. Article (CrossRef Link).
[15] Tomáš Pevný and Jessica Fridrich, "Towards multi-class blind steganalyzer for JPEG Images," in *Proc. of 4th Int. Workshop on Digital Watermarking*, *Spring Lecture Notes in Computer Science*, vol. 3710, pp. 39-53, Sepetember 15-17, 2005. Article (CrossRef Link).
[16] Jan Kodovský and Jessica Fridrich, "Calibration revisited," in *Proc. of 11th ACM Workshop on Multimedia and Security*, pp. 63-74, September 7-8, 2009. Article (CrossRef Link).
[17] Jing Dong, Wei Wang and Tieniu Tan, "Multi-class blind steganalysis based on image run-length analysis," in *Proc. of the 8th Int. Workshop on Digital Watermarking*, *Springer Lecture Notes in Computer Science*, vol. 5703, pp. 199-210, August 24-26, 2009. Article (CrossRef Link).
[18] Tomáš Pevný, Jessica Fridrich and Andrew D. Ker, "From blind to quantitative steganalysis," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 2, pp. 445-454, November, 2011. Article (CrossRef Link).
[19] Xiangyang Luo, Fenlin Liu, Chunfang Yang, Shiguo Lian and Daoshun Wang, "On F5 steganography in images," *The Computer Journal*, vol. 55, no. 4, pp. 447-456, April, 2012. Article (CrossRef Link).
[20] Zhenyu Li, Zongyun Hu, Xiangyang Luo and Bin Lu, "Embedding change rate estimation based on ensemble learning," in *Proc. of 1st ACM Workshop on Information Hiding Multimedia Security*, pp. 77-83, June 17-19, 2013. Article (CrossRef Link).
[21] Xiangyang Luo, Fenlin Liu, Shiguo Lian, Chunfang Yang and Stefanos Gritzalis, "On the typical statistic features for image blind steganalysis," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 7, pp. 1404-1422, August, 2011. Article (CrossRef Link).
[22] Ying Wang and Pierre Moulin, "Optimized feature extraction for learning-based image steganalysis," *IEEE Transactions on Information Forensics and Security*, vol. 2, no. 1, pp. 31-45, March, 2007. Article (CrossRef Link).
[23] Chih-Chung Chang and Chih-Jen Lin, "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 3, article 27, April, 2011. Article (CrossRef Link).

Jicang Lu received the B.S. degree and the M.S. degree and the Ph.D. degree from Zhengzhou Information Science and Technology Institute, Zhengzhou, China, in 2007, 2010 and 2014, respectively. From 2010 to 2011, he was a Visiting Scholar of the Department of Electric Engineering of Tsinghua University. Currently, he is a teacher of Zhengzhou Information Science and Technology Institute. His research interest includes image steganography and steganalysis technique.

Fenlin Liu received the B.S. degree from Zhengzhou Information Science and Technology Institute, Zhengzhou, China, in 1986, the M.S. degree from Harbin Institute of Technology, Harbin, Heilongjiang, in 1992, and the Ph.D. degree from the Northeast University, Shenyang, Liaoning, in 1998. He is currently a professor of Zhengzhou Information Science and Technology Institute. His research interests include information hiding and security theory.

Xiangyang Luo received the B.S. degree, the M.S. degree and the Ph.D. degree from Zhengzhou Information Science and Technology Institute, Zhengzhou, China, in 2001, 2004 and 2010, respectively. And his doctoral dissertation was awarded the National excellent doctoral dissertation nomination prize in 2012. From 2006 to 2007, he was a visiting scholar of the Department of Computer Science and Technology of Tsinghua University. He is currently an associate professor of Zhengzhou Information Science and Technology Institute. His research interest includes information security, image steganography and steganalysis.