Universal Image Steganalytic Method

Vladimír BÁNOCI, Martin BRODA, Gabriel BUGÁR, Dušan LEVICKÝ

Dept. of Electronics and Multimedia Communications, Technical University of Košice, Letná 9, 042 00 Košice, Slovak Republic

vladimir.banoci@tuke.sk, martin.broda@tuke.sk, gabriel.bugar@tuke.sk, dusan.levicky@tuke.sk

Abstract. *In the paper we introduce a new universal steganalytic method in JPEG file format that is detecting well-known and also newly developed steganographic methods. The steganalytic model is trained by MHF-DZ steganographic algorithm previously designed by the same authors. The calibration technique with the Feature Based Steganalysis (FBS) was employed in order to identify statistical changes caused by embedding a secret data into original image. The steganalyzer concept utilizes Support Vector Machine (SVM) classification for training a model that is later used by the same steganalyzer in order to identify between a clean (cover) and steganographic image. The aim of the paper was to analyze the variety in accuracy of detection results (ACR) while detecting testing steganographic algorithms as F5, Outguess, Model Based Steganography without deblocking, JP Hide&Seek which represent the generally used steganographic tools. The comparison of four feature vectors with different lengths FBS (22), FBS (66) FBS(274) and FBS(285) shows promising results of the proposed universal steganalytic method comparing to binary methods.*

Keywords

Steganography, universal steganalysis, message hiding, image processing, JPEG file format, statistical features.

1. Introduction

A recent development in digital data and extension of the Internet has promoted easier way of transferring of such data. Those networks serve enough bandwidth for data that possess high amount of redundancy, which can be exploited for implementation of subliminal communication. Therefore, sharing of multimedia is generally conveyed by two aspects. The first one is copyright infringement that can be resolved by watermarking or cryptographic methods. Secondly, multimedia data are also characterized by possessing of redundancy data that can be replaced or modified by steganographic tools in order to carry secret information without detecting to every-day user.

Steganographic methods can be used to illegal activities as terrorism or applied in environment where communication channel is monitored and user needs to transmit a secret message. This secret communication can be served by steganographic tools, where on the other hand, steganalysis is focused on detection and revealing of such communication mostly by employing a statistical analysis. Generally, the steganographic system is considered as broken once the steganalytic algorithm can decide whether testing medium contains a secret message or the original medium is intact while success probability is higher than random guessing.

In general, the image steganography utilizes LSB (Least Significant Bit) plane [1], [2] for embedding a secret message or modifies coefficients of selected discrete transformation, e.g. Discrete Cosine Transform (DCT). The LSB substitution is considered as one of the first methods that replaces LSB or any lower planes of the image pixels with the bits of a secret message. In these methods, the pixels of the cover image are chosen either sequentially or randomly. The variety of techniques including spread spectrum steganography [3], statistical steganography [4] and generation-based steganography [5] can be combined with the general embedding concept that was in later research also adjusted to statistical steganalysis. Some of these techniques are complemented by game theory analysis [6] and cryptography [7] in order to obstruct a successful steganalysis and extraction a relevant statistical data.

The steganalytic methods in static images are divided basically into two categories. The first one is binary steganalysis where steganalytic methods are primarily designed for detection of specific steganographic algorithms as F5 [8], Outguess [9], MB [10], JPHS [11] and others. The second category represents universal steganalysis methods whose steganalyzers are able to detect various steganographic algorithms. In addition, the universal steganalysis is designed to detect newly developed steganographic methods. However, the detection efficiency of universal steganalytic methods is significantly lower for methods that belong to specific steganalysis.

The process of universal steganalysis consists of two essential steps: the feature extraction needed for construction of the model and testing of the image by a steganalyzer with a selected model. The first part relies on unambiguously defined image sets consisting of cover and stego images that are used for feature extraction. The aim is to train a model based on classification techniques like SVM.

After reconstructing a classifier's model based on calculated features, the classifier is able to decide whether the testing image possesses a secret message embedded by a steganographic method.

The recent steganalytic methods are mostly based on three principles according to acquiring of statistical features. The first category relies on Binary Similarity Measures (BSM) wherein e.g. Avcibas [12] proposed a steganalytic method without using a reference image and processing only 18 parameters. The basic idea arises from a correlation between the bit planes as well as the binary texture characteristics within bit planes which differ between stego and cover image. Some methods in this category extract features in spatial domain, while the others directly in transform domain from coefficients [13].

The second approach of steganalytic method uses a model trained by first and higher-order wavelet statistics that are able to distinguish between images with and without hidden messages. In the proposed methods [14], [15], the features are extracted from image decomposed levels in wavelet domain. Hence, the transform areas that are sensitive to embedded information are used in extraction of statistical features.

Finally, the third popular steganalytic principle is based on statistical analysis of the first and second order statistics in transform domain in images (Feature Based Steganalysis (FBS)). Those methods are differentiated according to extraction process of statistical data and the approach of calculating those statistics. At the beginning, the first attempts relied on first order statistics as 'chi-square attack' proposed by Westfield [16] and this method was later improved with a random message embedding [17]. Another approach applied directly in JPEG on transform coefficients was proposed by Fridrich [18], whose universal steganalytic method utilizes an image calibration. The calibration crops stego image by 4 pixels in each direction in order to suppress spatial blockiness effect that appears during embedding a secret message. After calibration, the resulting image has characteristics very similar to the original image. The calculated difference between the calibrated and original image represents a statistical feature. Consequently, a retrieved feature vector (with defined length) represents an input for steganalytic classifier.

Our proposed universal steganalytic method utilizes a feature extraction in DCT domain in JPEG images [19] using a model that was trained on images created by MHF-DZ steganographic method. This paper is organized as follows. In the following section a general scheme of the novel steganographic method is explained that is used for training of steganalytic model. The next section describes specific steganalytic algorithm based on feature extraction in JPEG images. In Section 4 experimental results of detection accuracy for the proposed universal steganalytic model are given also with comparison of binary classification models. Experiments of detection accuracy were carried out for popular steganographic tools that are used nowadays.

2. MHF-DZ as a Blind Steganographic Method

The main steganography objective is to ensure a secret transmission of data in background of non-privacy communication without additional security elements, e.g. cryptographic keys. Security in this case is based on confidentiality of transmission. Detection of secret transmission exposes communication security itself what could result in disclosure of secret message to undesirable party of communication.

Steganographic methods can be classified into two categories according to extraction process of secret message from stego data. The first category consists of steganographic methods which require also original data for successful extraction of a secret message, whereas methods in the second category do not require additional information, i.e. for extraction of a secret message only stego data and knowledge of extraction algorithm are necessary. Algorithms in this category are also called as blind steganographic systems.

The steganographic method MHF-DZ (Modulo Histogram Fitting with Dead-Zone) is derived from MHF-AES method, which the authors recall in [20]. The algorithm uses modulo arithmetic for embedding a secret message in JPEG file format. Embedding of secret message is performed by alternation of DCT coefficients in such a manner that calculated coefficient's modulo equals to a secret symbol. The modifying value of DCT coefficients is given by modulo window (MW) that induces an ambiguity about addressing of modifying value for specific reasons being deeply discussed in the paper. An additional tracking algorithm addresses the best-fit modifying value according to the frequency component's histogram of cover image. The provided algorithm allows compensation of changes in initial histogram by analyzing of previously embedding changes in DCT coefficients.

Application of modulo arithmetic gives predisposition for a blind steganographic system which does not require a presence of a cover-image on the receiver side. The method's capacity is not restrained, however, using higher modulo $n > 2$ makes the method vulnerable against histogram analysis and also providing a trade-off between quality of stego image and transfer capacity. The general scheme of MHF-DZ algorithm is shown in Fig. 1.

Implementation of AES ciphering block was introduced in order to get uniform probability distribution of a secret message symbols which is not adequate mapping to DCT coefficient probability distribution that is assigned to m-state quantizer. However, integration ciphering block improves after-embedding histograms of DCT frequency components. If distribution conformity between embedding symbols and transformation coefficients is not fulfilled, the after-embedding histogram shows significant distortion in form of "tooth-run" character. The zero value coefficients are defined by predominant occurrence among all transformation coefficients due to application of JPEG quanti-

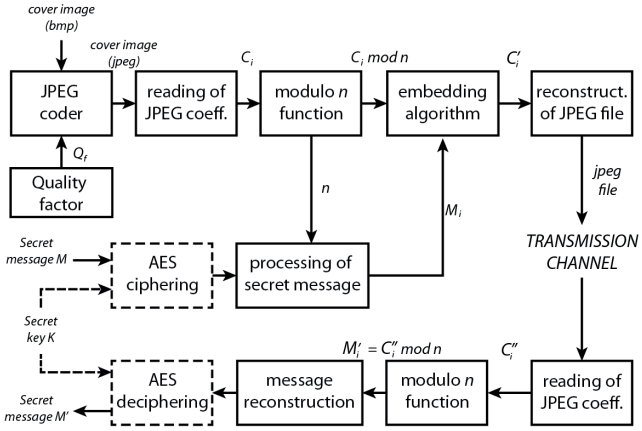

Fig. 1. General scheme of MHF-DZ steganographic system.

zation table. In respect to previous statement, the above-mentioned method was complemented with *Dead-Zone (DZ)* area which defines a number of zero-value coefficients (for each frequency component *AC*) that are not used in embedding process in order to comply the probability distribution for $n = 2$. However, the *DZ* information needs also be transferred altogether with the secret message for successful reconstruction of the secret message.

Moreover, the previously designed method MHF-DZ based on histogram preserving scheme proves the resistance against proposed statistical steganalyzer for low embedded rate of secret data. The higher embedding rate ($bpnz \geq 0.5$) involves changes for more zero DCT coefficients of higher frequencies into embedding process what results in a detectable distortion of local histograms as other intra or inter block corruptions of DCT coefficients.

3. Feature Based Steganalysis

If the process of steganalysis is able to reveal secret communication, steganographic system is defined as broken and the purpose of steganography is marred. Steganalytic method is defined as successful when stego image can be differentiated from cover image with higher probability than random guessing. Steganalysis can be supplemented by activity of extraction secret message intelligence what requires a set of techniques for further analysis and increase of computational demands [21]. The universal (blind) steganalytic methods are defined as sets of detection techniques that are independent to the applied algorithm and achieve good detection results of embedded message that was hidden by new or unknown steganographic methods. If steganalytic technique is adjusted to steganographic method and its characteristics then this technique can achieve higher efficiency in the process of detection. This type of steganalysis is referred as special or binary steganalysis.

The main objective of steganalysis in static images is detection of changes in statistic properties of cover image due to embedding a secret message. Therefore, the calculation of those statistical features is crucial in design of steg-

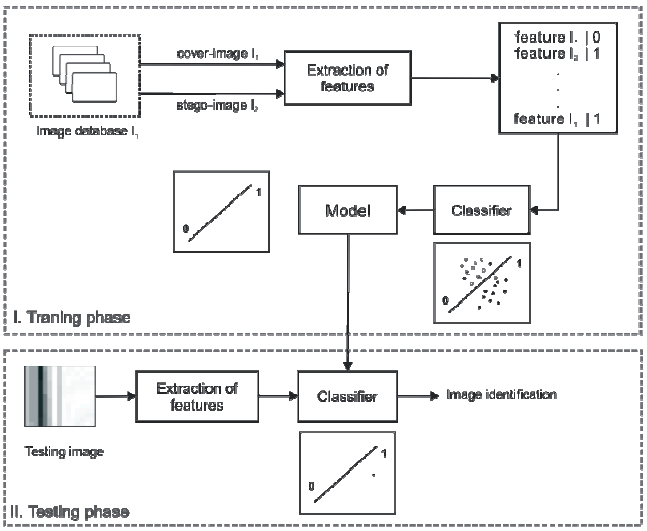

Fig. 2. Basic diagram of Feature Based Steganalysis.

analytic method. The obtained statistical features represent the input for classifier block in training phase as it is illustrated in Fig. 2.

The Feature Based Steganalysis (FBS) was implemented altogether with calibration technique proposed by Fridrich [22] that performs cropping of picture by 4 pixels in each direction. The calibrated image has very similar statistical features to the cover image. The calibration of statistical parameters is also important in feature extraction. The block diagram of feature extraction with image calibration is illustrated in Fig. 3.

Fig. 3. Feature Based Steganalysis with image calibration.

The image database consists of several thousand of images that were taken by different types of cameras using different camera's settings and resolutions. Stego images are created by embedding a secret message with several steganographic methods (e.g. F5, Outguess, MB and others used in JPEG files). In the next step, statistical features are extracted from stego or cover images, whereby we obtain two sets of statistical parameters that are separated according to identifier.

In our case, we define four different lengths of statistic vector:

- FBS (22) – 22 statistical parameters,
- FBS (66) – 66 statistical parameters,
- FBS (274) – 274 statistical parameters,
- FBS (285) – 285 statistical parameters.

The first group of statistical features is denoted as FBS (22) which consists of the global histogram (1) [21], [24] from all $64 \times n_B$ DCT (where B is the number of total 8×8 blocks in the image) and the local histogram (2) in

mode $(i, j) \in (2,1)$ [21], [24]. The central part <-5,5> of this histogram was selected due to the situated maximum of the global histogram.

$$H = (H_L,....,H_R), \tag{1}$$

$$h^{i,j} = (h_L^{i,j},...,h_R^{i,j}). \tag{2}$$

The second group of statistical parameters FBS (66) involves parameters as global (1) and local histograms (2) in modes $(i, j) \in \{(1,2),(2,1),(3,1),(2,2),(1,3)\}$.

The third group FBS (274) contains all mentioned FBS (66) statistical parameters with additional parameters, which are defined as follows [23], [24]:

Dual histograms (3) for matrix 8×8 $g_{i,j}^{d}$ for values i, j = 1,…,8 and d = -5,…,5 compose next 11 functions.

$$g_{i,j}^{d} = \sum_{k=1}^{n_B} \delta(d, d_{i,j}(k)) \tag{3}$$

where $\delta(x, y) = 1$, if $x = y$ and $\delta(x, y) = 0$, if $x \neq y$.

Next functions are acquired from intra-blocking dependencies of DCT coefficients. First of them is variation (4).

$$V = \frac{\sum_{i,j=1}^{8}\sum_{k=1}^{|I_r|-1}\left|d_{i,j}(I_r(k))\right| - \left|d_{i,j}(I_r(k+1))\right|}{|I_r| + |I_c|} + \frac{\sum_{i,j=1}^{8}\sum_{k=1}^{|I_c|-1}\left|d_{i,j}(I_c(k))\right| - \left|d_{i,j}(I_c(k+1))\right|}{|I_r| + |I_c|} \tag{4}$$

where I_r and I_c represent vectors of block that is traversing in horizontal (r) and vertical (c) direction in the image and k = 1, ..., n_B.

Integral measure of intra-blocking dependence (5) is calculated from spatial representation of JPEG image.

$$B = \frac{\sum_{i=1}^{\lfloor (M-1)/8 \rfloor}\sum_{j=1}^{N}\left|c_{8i,j} - c_{8i+1,j}\right|^{\alpha} + \sum_{j=1}^{\lfloor (N-1)/8 \rfloor}\sum_{i=1}^{M}\left|c_{i,8j} - c_{i,8j+1}\right|^{\alpha}}{N\lfloor (M-1)/8 \rfloor + M\lfloor (N-1)/8 \rfloor} \tag{5}$$

where α = 1,2 and M, N are dimension of image and C_{ij} represents the luminance component of uncompressed JPEG image. Other functions (6), (7), (8) are calculated from co-occurrence matrix **C** of neighboring DCT coefficients for image J_1 (stego image) and J_2 (image after calibration itself crops stego-image by 4 pixels in each direction). Matrix **C** describes the probability distribution of pairs of neighboring DCT coefficients.

$$N_{00} = C_{0,0}(J_1) - C_{0,0}(J_2), \tag{6}$$

$$N_{01} = C_{0,1}(J_1) - C_{0,1}(J_2) + C_{1,0}(J_1) - C_{1,0}(J_2) + C_{-1,0}(J_1) - C_{-1,0}(J_2) + C_{0,-1}(J_1) - C_{0,-1}(J_2), \tag{7}$$

$$N_{11} = C_{1,1}(J_1) - C_{1,1}(J_2) + C_{1,-1}(J_1) - C_{1,-1}(J_2) + C_{-1,1}(J_1) - C_{-1,1}(J_2) + C_{-1,-1}(J_1) - C_{-1,-1}(J_2). \tag{8}$$

The last mentioned functions represent the Markov model. The Markov model defines diverseness between absolute values of adjacent DCT coefficients. An embedding with steganographic methods entails defects in smoothness, regularity, continuity and consistence what means that steganographic methods can corrupt correlation between DCT coefficients in cover image.

The Markov model is defined by the following consequent process. In the first step, absolute values $F(u,v)$ of coefficients are calculated. These coefficients are collocated to the same process as picture elements in the original image. From absolute values differential arrays in horizontal, vertical, diagonal direction and differential array with minor diagonal ($F_h(u,v)$, $F_v(u,v)$, $F_d(u,v)$, $F_m(u,v)$) are calculated. These different arrays can be modeled using Markov processes, whereby the transition probability matrixes $\boldsymbol{M}_h$ (9), $\boldsymbol{M}_v$ (10), $\mathbf{M}_d$ (11), $\mathbf{M}_m$ (12) are acquired as follows:

$$M_h(i,j) = \frac{\sum_{u=1}^{S_u-2}\sum_{v=1}^{S_v}\delta(F_h(u,v) = i, F_h(u+1,v) = j)}{\sum_{u=1}^{S_u-1}\sum_{v=1}^{S_v}\delta(F_h(u,v) = i)}, \tag{9}$$

$$M_v(i,j) = \frac{\sum_{u=1}^{S_u}\sum_{v=1}^{S_v-2}\delta(F_v(u,v) = i, F_v(u,v+1) = j)}{\sum_{u=1}^{S_u}\sum_{v=1}^{S_v-1}\delta(F_v(u,v) = i)}, \tag{10}$$

$$M_d(i,j) = \frac{\sum_{u=1}^{S_u-2}\sum_{v=1}^{S_v-2}\delta(F_d(u,v) = i, F_d(u+1,v+1) = j)}{\sum_{u=1}^{S_u-1}\sum_{v=1}^{S_v-1}\delta(F_d(u,v) = i)}, \tag{11}$$

$$M_m(i,j) = \frac{\sum_{u=1}^{S_u-2}\sum_{v=1}^{S_v-2}\delta(F_m(u+1,v) = i, F_m(u+1,v) = j)}{\sum_{u=1}^{S_u-1}\sum_{v=1}^{S_v-1}\delta(F_m(u,v) = i)} \tag{12}$$

where S_u and S_v represents picture dimension. Generally, the Markov model increases sensitivity of specific and universal steganalysis methods in detection of a secret communication. However, in some cases, when the Markov model is not involved in FBS, the performance of steganalyzer demonstrates better results as it will be presented in the next section.

The last proposed group FBS(285) includes all the previous statistical parameters mentioned in group FBS(274) together with statistics based on inter blocking dependence. DCT image coefficients are divided into matrices with size 64 × n_B and consequently the difference between adjacent blocks of DCT coefficients on equivalent positions is calculated (13).

$$D_{i,j} = d_{i,j} - d_{i,j+1} \tag{13}$$

where $D_{i,j}$ is matrix that is calculated by the difference between all adjacent blocks of DCT coefficients using horizontal sampling. Consequently histogram (14) is defined from this matrix in interval <-5, 5> where its maximum is situated.

$$D = (D_L, ..., D_R) \tag{14}$$

where $L = \min_{i,j,k} d_{i,j}$ a $R = \max_{i,j,k} d_{i,j}$.

By this approach we can get additional 11 statistical parameters, whereby we obtain the statistical vector with the length 285.

The next block in the steganalytic scheme is a classifier. The classifier's input is represented by a set of statistical features of stego and cover images. The result of classifier process is a trained model grouping cover images from stego images that were obtained by specific steganographic method. In this case, we are denoting this process as binary classification. Support Vector Machine (SVM) classifier [25] was introduced in tested feature based steganalytic method. For example, output binary classifier is model between *Cover images – F5 stego images, Cover images – Outguess stego images* and other.

4. Experimental Results

In training process, the image database contains stego and cover images in JPEG file format. Cover images were taken by different types of camera and with various image resolutions (1400 × 1000, 1270 × 852, 960 × 720, 800 × 600, 720 × 540 and 640 × 480). Stego image were obtained by F5, Outguess, Model Based (MB), JPEG Hide&Seek (JPHS) and MHF-DZ steganographic methods. After embedding, the final database consists of 24000 images for feature extraction and consequently processing of the classifier. A part of feature extraction is calibration technique that was explained in the previous section. The calibration was executed on image database in order to acquire difference statistics of DCT coefficients what means a feature vector. In training phase, steganalytic models for binary classification were created, e.g. model *cover – F5 stego images, cover - Outguess stego images* etc. for each tested steganographic method.

In testing phase, two types of experiments were realized. The first test verifies detection accuracy of created models for specific steganographic methods - binary classification. In the second part of experiment, the model that was created from statistic features of cover and stego images using MHF-DZ steganographic tool was tested. The model trained by the same tool was later used for classification of images that were created by other steganography tools (F5, Outguess, MB, and JPHS) in order to demonstrate universal steganalysis detectability.

The steganalyzer performance is highly susceptible to embedded data rate that is given by *bpnz (bit per non-zero)* DCT coefficients and also to length of feature vector. The tested steganographic methods exhibit non-equal embedding capacity which does not make it possible to show comparable results of final detection Accuracy (ACR) and True Positive Rate (TPR) for all given values of *bpnz*.

Testing algorithm	bpnz	FBS (22)		FBS (66)		FBS (274)		FBS (285)	
		TPR[%]	ACR[%]	TPR[%]	ACR[%]	TPR[%]	ACR[%]	TPR[%]	ACR[%]
F5	0.25	58.4	78.2	82.8	89.6	36.8	68.4	49.1	75.1
	0.4	89.2	93.6	96.8	96.6	73.2	86.6	76.4	88.7
	0.6	98.4	98.2	100	98.2	94	97	95.1	97.5
	0.8	100	99	100	98.2	96.8	98.4	97.5	98.9
Outguess	0.05	59.6	65.4	70	71.2	53.6	74.6	56.9	76.1
	0.1	70.8	71	89.6	81	89.6	92.6	90.1	94.1
	0.15	84	77.6	96.4	84.4	95.6	95.6	95.1	95.9
	0.2	84.8	78	98.4	85.4	98.4	97	98.5	97.9
	0.25	90.3	80.7	97.9	85.2	98.8	97.2	98.9	98.1
	0.4	89.5	80.4	98.2	85.3	100	97.8	100	98.1
MHF-DZ	0.2	7.3	39.3	23.8	48.1	15.2	55.4	22.1	58.1
	0.4	9.2	53.6	45.8	66.9	38.3	66.8	40.1	69.1
	0.6	21.1	59.6	64.4	76.2	65.6	80.4	65.7	80.9
	0.8	38	68	73.2	80.6	77.5	86.4	78.7	87.6
	1	56	77	74	81	90	92.6	91.3	93.4
MB	0.1	47.6	48.4	54.8	50.8	70.4	83.8	72.6	85.4
	0.2	41.6	45.4	58	52.4	99.6	98.4	100	98.9
	0.4	29.2	39.2	56.4	51.6	100	98.6	100	98.9
	0.6	18	33.6	52.8	49.8	100	98.6	100	98.9
	0.8	8.4	28.8	47.6	47.2	100	98.6	100	98.9
	1	4	26.6	42	44.4	100	98.6	100	98.9
JPHS	0.1	67.6	65.4	84.2	69.1	72.9	74.5	73.1	76.1
	0.25	86.1	74.7	94.7	74.4	98.4	87.2	98.4	89.1
	0.4	90.2	76.7	96.2	75.1	99.1	87.6	99.3	89.4
	0.5	97.3	80.3	98.7	76.4	100	88	100	90.1
	0.6	98.8	81	100	77	100	88	100	90.3

Tab. 1. True Positive Rate and Accuracy of detection for binary classification of specific steganographic methods.

Tab. 1 shows TPR/ACR values with various embedding rate of *bpnz* ranging from the interval where the steganographic method is still secure, and contrary where the method has proven as successfully detectable by steganalyzer or the method's capacity is reached. The results show that in case of F5 steganographic method, the binary steganalyzer's performance acquires the best detection results with feature length FBS(66) and FBS(22) where the higher length of feature vector FBS(285) does not improve steganalyzer's susceptibility to detecting the embedding method. Nevertheless, for other tested methods, the steganalyzer's performance rises along with higher length of calculated features where the best result is accounted for FBS(285) that means also including Markov model and features of inter-blocking dependence.

Comparing with other research publications in steganalysis e.g. [26] that use similar FBS steganalysis approach but different classifier (J48, SMO, Naive Bayes), our proposed system accounts for better results in detection of F5 and JPHS method. The best result for detection of F5 method with embedding load *bpnz* = 0.25 was achieved with SMO classifier and FBS(246) feature set at the level of ACR = 85.7 % what is lower than our result ACR = 89.7 % with SVM classifier and FBS(66) feature set. Diversion of detection results for JPHS methods is even more distinct where ACR = 57.15 % (*bpnz* = 1) is achieved with SMO and FBS(246) comparing to our ACR = 100 with SVM and FBS (66) even with lower *bpnz* = 0.6.

Figure 4 shows graphic representation of detection accuracy of the tested steganographic methods for 285 statistical features and *bpnz* = 0.4.

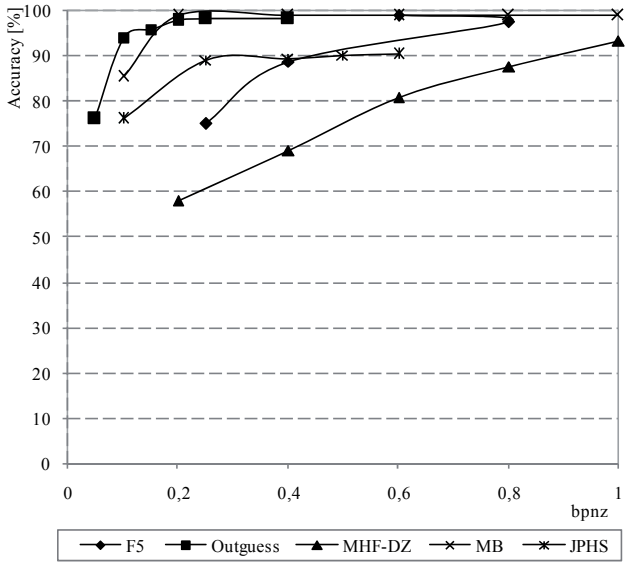

Fig. 4. ACR of binary classification of steganographic methods for FBS (285).

Detection accuracy for steganalytic model can be illustrated using ROC (Receiver Operating Characteristic) curve. Figure 5 illustrates ROC curves for specific models of steganographic methods with FBS(285) and *bpnz* = 0.4.

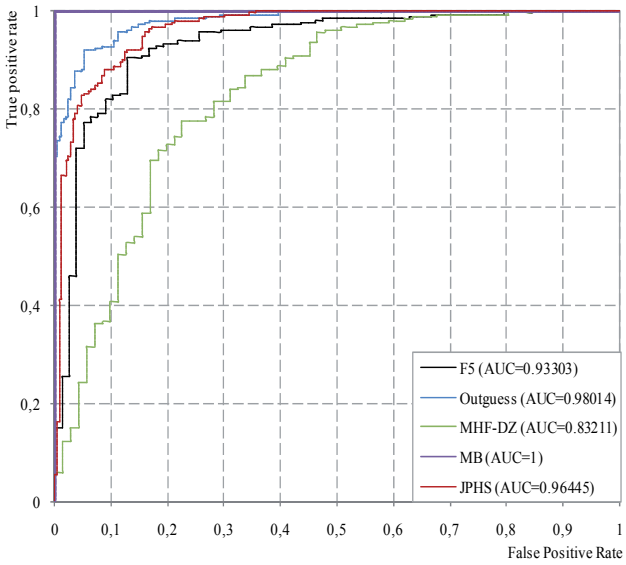

Fig. 5. ROC curves of binary classification of tested steganographic methods with FBS (285) and *bpnz*=0.4.

Many of steganographic methods are trying to preserve global and local histograms of DCT coefficients of cover image as it is present in MHF-DZ method. The difference between features of calibrated image and stego image rises proportionally with a number of tested histograms and thus increases the risk of detection by steganalyzer.

The results of universal steganalyzer based on MHF-DZ method are shown in Tab. 2 where ACR results of universal classification are lower comparing to binary classifi

Testing alghorithm	bpnz	Trained model MHF-DZ / Cover ACR [%]			
		FBS (22)	FBS (66)	FBS (274)	FBS (285)
F5	0.25	78.8	69.2	68.4	72.1
	0.4	92.4	75.4	77.2	80.1
	0.6	98.6	82	86.4	88.5
	0.8	99	87	90.8	92.1
Outguess	0.05	66	64.2	64.4	66.3
	0.1	67.6	64.4	65.2	67.4
	0.15	69.2	65.2	64.8	66.1
	0.2	70.2	66.4	66.4	68.7
	0.25	77.1	71.5	74.4	75.9
	0.4	68.5	59.8	64	65.3
MHF-DZ	0.2	39.3	48.1	55.4	58.1
	0.4	53.6	66.9	66.8	69.1
	0.6	59.6	76.2	80.4	80.9
	0.8	68	80.6	86.4	87.6
	1	77	81	92.6	93.4
MB	0.2	67	79.6	55.6	57.1
	0.4	90.4	91.8	69	71.2
	0.6	96	94	83	84.6
	0.8	97.4	94	90.8	92.1
	1	98.2	94	94.4	94.5
JPHS	0.25	52.5	50.8	67.6	68.9
	0.4	54.3	62	73.8	75.4
	0.5	56.4	52.7	80.8	81.9
	0.6	59.2	53.6	89.2	89.8

Tab. 2. Accuracy of universal classification of tested steganographic methods based on MHF-DZ model.

cation ranging from ~4 % to ~30 % in relation to feature set. The detection rate for F5 method is degraded by 4.3 %, Outguess by 30 %, MB method by 6.9 % and JPHS method by 15.7 %. The results are achieved with different sets of features, whereas e.g. F5 method, in case of binary classification, allocates the best detection results using FBS(66) feature set and for universal classification the highest detection rate is achieved by FBS(22) feature set. Moreover, interesting ACR results with 230 % increase were achieved for MB algorithm using FBS(22) feature set. The same feature set was used for detection of F5 method, where just 4.3 % decrease of ACR was allocated comparing to binary classification. Generally, the advantage of FBS(22) against other feature sets of higher length could be seen in faster calculating of statistics in contrary to FBS(285).

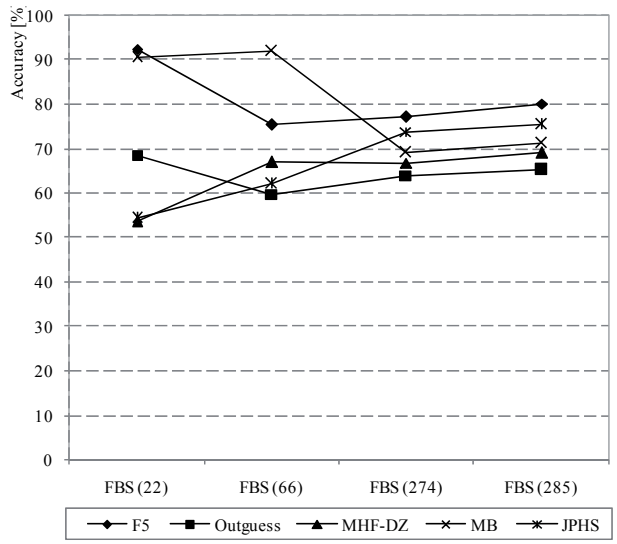

Fig. 6. ACR rate of universal classification for different length of FBS and *bpnz*= 0.4.

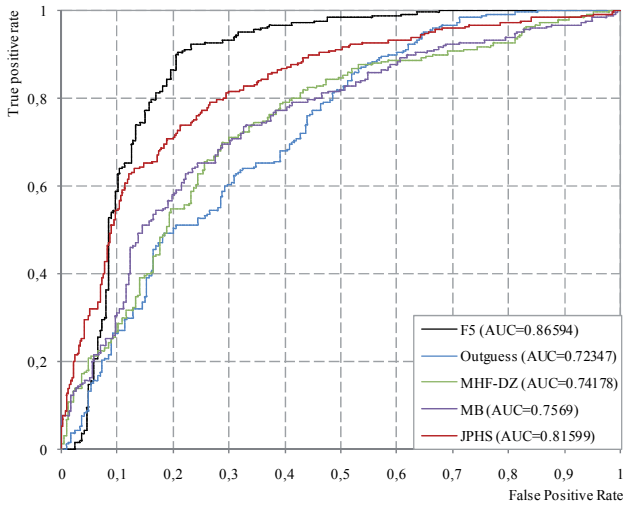

Fig. 7. ROC curves of specific steganographic methods for FBS (285) and *bpnz* = 0.4

Figure 6 illustrates ACR rates of universal classification in order to show the best attained results based on different FBS sets. Achieved ACR = 53% result for JPHS shows that detection using universal model (MHD-DZ / Cover) and FBS(22) set is not satisfactory for this method. Hence, higher length of feature set needs to be employed.

Figure 7 illustrates ROC curves for detection of tested steganographic tools applying the proposed universal steganalysis based on MHF-DZ method.

5. Conclusion

In this paper, we proposed universal steganalytic method in still images in JPEG file format for detection secret communication using several steganographic algorithms. The aim of our research was to introduce a method that is able to detect well-known methods (F5, Outguess, MB, JPHS) as well as a newly developed steganographic methods with satisfactory detection rate. The comparative results with binary classification were provided to evaluate detection performance of the universal classifier and its model. The ACR results of universal classifier are lower as expected comparing to binary classification. However, the steganalyzer's usage is not restrained to only one steganographic tool but is capable to detect multiple steganographic algorithms. The trained model is based on MHF-DZ steganographic tool that is masking its presence during embedding with histogram fitting scheme of DCT coefficients. The proposed method uses FBS approach for extraction statistical features and SVM for its classification. The feature vector represents the statistical difference between stego and cover image, where its length defines the quantity of analyzed statistics. The same model with four various feature statistics was tested (FBS(22), FBS(66), FBS(274) and FBS(285)). The results show that for some methods (MB, F5), features with lower length FBS(22) show significantly better detection results comparing to high order statistics testing what is in contrary to binary classification scenario, where FBS(285) accounts for better ACR rate results in general. Nevertheless, due to high number of false detection for cover images by Outguess model, utilization of FBS(285) is also recommended for our proposed method.

Acknowledgements

The paper is the result of the Project implementation: Ministry of Education of the Slovak Republic VEGA Grant No. 1/0386/12 and University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDF.

We support research activities in Slovakia/This project is being co-financed by the European Union.

References

[1] SHARP, T. An implementation of key-based digital signal steganography. In *4th International Workshop Information Hiding. Lecture Notes in Computer Science*, 2001, vol. 2137, p. 13–26.

[2] JOHNSON, N. F., JAJODIA, S. Exploring steganography: Seeing the unseen. *IEEE Computer*, 1998, vol. 31, no. 2, p. 26–34.

[3] MARVEL, L., BONCELET, C., RETTER, C. Spread spectrum image steganography. *IEEE Transactions on Image Processing*, 1999, vol. 8, no. 8, p. 1075–1083.

[4] SARKAR, A., MANJUNATH, B. S. Estimating steganographic capacity for odd even based embedding and its use in individual compensation. In *Proc. IEEE Int. Conf. on Image Processing (ICIP2007*). San Antonio (TX, USA), 2007, p. 409–412.

[5] KATZENBEISSER, S., PETITCOLAS, F. A. P. *Information Hiding Techniques for Steganography and Digital Watermarking*. Artech House Books, 2000.

[6] KER, D. A. Batch steganography and the threshold game. In *Proc. SPIE Conference Vol. 6505, Security, Steganography, and Watermarking of Multimedia Contents IX*. 2007, p. 0401–0413. DOI: 10.1117/12.703334.

[7] MARWAHA, P., MARWAHA, P. Visual cryptographic steganography in images. In *2nd Internat. Conf. on Computing Communication and Networking Technologies*. Karur (India), 2010, p. 1–6.

[8] WESTFELD, A. High capacity despite better steganalysis (F5 – a steganographic algorithm). In *4th Internat. Workshop Information Hiding. Lecture Notes in Computer Science.* 2001, vol. 2137, p. 289–302.

[9] PROVOS, N. Defending against statistical steganalysis. In *Proc. 10th USENIX Security Symposium*. Washington (USA), 2001, p. 323–335.

[10] SALLEE, P. Model-based methods for steganography and steganalysis. *International Journal of Image and Graphics*, 2005, vol. 5, no. 1, p. 167–190.

[11] HOPPER, N. J., LANGFORD, J., VON AHN, L. Provably secure steganography. In YUNG, M. (ed.) *Advances in Cryptology CRYPTO 2002. LNCS. 22nd Annual International Cryptology Conference*. Heidelberg: Springer, 2002, vol. 2442, p. 77–92.

[12] AVCIBAS, I., KHARAZZI, M., MEMON, N., SANKUR, B. Image steganalysis with binary similarity measures. *EURASIP J. Applied Signal Processing*, 2005, no. 17, p. 2749–2757.

[13] LIN, J-Q., ZHONG, S-P. JPEG image steganalysis method based on binary similarity measures. In *Proc. of the 8th Internat. Conf. on Machine Learning and Cybernetics*. Baoding (China), 2009, p. 2238–2241.

[14] FARID, H., LYU, S. Higher-order wavelet statistics and their application to digital forensics. In *IEEE Workshop on Statistical Analysis in Computer Vision (in conjunction with CVPR)*. Madison, 2003, p. 1-8.

[15] ZONG, H., LIU, F., LUO, X. Blind image steganalysis based on wavelet coefficient correlation. *Digital Investigation*, 2012, vol. 9, no. 1, p. 58–68.

[16] WESTFELD, A., PFITZMANN, A. Attacks on steganographic systems. In *3rd Internat. Workshop Information Hiding'99. Lecture Notes in Computer Science,* 2000, vol. 1768, p. 61–76.

[17] PROVOS, N., HONEYMAN, P. Detecting steganographic content on the Internet. *CITI Technical Report* 01-11, 2001.

[18] FRIDRICH, J. Feature-based steganalysis for JPEG images and its implications for future design of steganographic schemes. In *6th Internat. Workshop on Information Hiding. Lecture Notes in Computer Science*, 2005, vol. 3200, p 67–81.

[19] MAJERČÁK, D., BÁNOCI, V., BRODA, M., BUGÁR, G., LEVICKÝ, D. Performance evaluation of feature-based steganalysis in steganography. In *23rd International Conference Radioelektronika*. Pardubice (Czech Republic), 2013, p. 377–382.

[20] BÁNOCI, V., BUGÁR, G., LEVICKÝ, D., KLENOVIČOVÁ, Z. Histogram secure steganography system in JPEG file based on modulus function. In *22nd International Conference Radioelektronika 2012*. Brno (CR), 2012, p. 263–266.

[21] BOHME, R. *Advanced Statistical Steganalysis*. Dresden: Springer, 2009. ISBN 978-3-642-14312-0

[22] FRIDRICH, J., GOLJAN, M., HOGEA, D. Steganalysis of JPEG images: Breaking the F5 algorithm. In *5th Internat. Workshop Information Hiding. Lecture Notes in Computer Science*, 2003, vol. 2578, p. 310–323.

[23] FARID, H., SIWEII, L. Detecting hidden messages using higher-order statistics and support vector machines. In *5th Internat.l Workshop Information Hiding. Lecture Notes in Computer Science*, 2003, vol. 2578, p. 340–354.

[24] FRIDRICH, J., PEVNÝ, T. Merging Markov and DCT Feature for Multi-Class JPEG Steganalysis. In *Proc. SPIE Conference Vol. 6505, Security, Steganography, and Watermarking of Multimedia Contents IX*. 2007. DOI: 10.1117/12.696774

[25] CHANG, C-C, LIN, C-J. *LIBSVM: A library for support vector machines.* Software available at http://www.csie.ntu.edu.tw/cjlin/libsvm, 2001

[26] ASHU, CHHIKARA, R. Performance evaluation of first and second order features for steganalysis. *International Journal of Computer Applications,* April 2014, vol. 92, no. 16, p. 17–22.

About Authors ...

Vladimír BÁNOCI was born in Košice, Slovak Republic. He received his M.Sc. and PhD. degree from the Faculty of Electrical Engineering and Informatics, Technical University in Košice and now he continues as post-doc researcher at the same faculty. His research interests include image processing, spread spectrum communication, watermarking, steganography and steganalysis.

Martin BRODA was born in Prešov, Slovak Republic in 1988. He received his M.Sc. degree from the Faculty of Electrical Engineering and Informatics, Technical University in Košice. Nowadays, he is a PhD. student at the Dept. of Electronics and Multimedia Communications, focusing on multimedia security, image steganography, steganalysis and digital watermarking.

Gabriel BUGÁR was born in Košice. He received his M.Sc. and PhD. degree from the Faculty of Electrical Engineering and Informatics, Technical University in Košice and now he is currently holding a position of assistant professor. His research interests include image processing, steganography and steganalysis.

Dušan LEVICKÝ was born in Slanec, Slovak Republic in 1948. He received his M.Sc. and PhD. degrees at the Technical University (TU) in Košice and now he is a professor at the Dept. of Electronics and Multimedia Communications, TU in Košice. His research interests include digital image processing, image transmission and cryptography.