This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TDSC.2019.2962672, IEEE Transactions on Dependable and Secure Computing

Adaptive Steganalysis Based on Statistical Model of Quantized DCT Coefficients for JPEG Images

Tong Qiao, Xiangyang Luo, Ting Wu, Ming Xu and Zhenxing Qian

Abstract—In current steganalysis, relying on a large scale of samples, widely-adopted supervised schemes require the training stage while few studies focus on the design of a training-free unsupervised adaptive detector with high efficiency. To fill the gap, we investigate an adaptive statistical model-based detector designed for detecting JPEG steganography. First, in virtue of hypothesis testing theory, together with the distribution of quantized DCT coefficients, we establish the general framework of the statistical model-based detector. Second, based on the framework, we mainly analyze the performance of the detector relying on the selection of the statistical model, parameters estimation, and less significant payload prediction. Third, to improve the reliability of detection, based on the strategy of assigning weights for DCT channels, the novel adaptive statistical model-based detectors are proposed to aim at detecting JPEG steganography, involving the channel-selected or non-channel-selected algorithm. Extensive experiments highlight the effectiveness of the proposed methodology. Moreover, when detecting JPEG images adopted by two steganographic schemes with the small payload, the experimental results show the Area Under Curve (AUC) of our proposed optimal adaptive detector can achieve as high as 0.9567 and 0.9895 respectively, which are both better than that of non-adaptive detector.

Index Terms—JPEG steganalysis, statistical model-based detector, hypothesis testing, distribution model, DCT coefficients selection.

1 Introduction

Steganography focuses on embedding a secret message into cover media. Conversely, steganalysis aims to identify stego media containing hidden information. In general, the hidden bits are embedded in the spatial domain or frequency domain such as Discrete Cosine Transform (DCT)[1] coefficients. By considering the undetectability, current image steganography prefers embedding hidden bits in the regions which are difficult to statistically model, also namely adaptive steganography. Correspondingly, it makes sense that the investigation of adaptive steganalysis remains a hot topic.

The problem of image steganalysis has been addressed for many years. However, most detectors are designed via the supervised learning-based strategy while few works focus on training-free scheme. Moreover, the detection performance of designed supervised detectors mainly relies on the varieties of training samples. Recently, the training-free statistical model-based detectors have been advanced while they cannot completely address the new challenge from adaptive steganography. Therefore, in this paper, we propose to design adaptive statistical model-based detectors.

Tong Qiao is with School of Cyberspace, Hangzhou Dianzi University, Hangzhou, China; State Key Laboratory of Mathematical Engineering and Advanced Computing, Zhengzhou Science and Technology Institute, Zhengzhou, China; email: tong.qiao@hdu.edu.cn
Xiangyang Luo (corresponding author) is with State Key Laboratory of Mathematical Engineering and Advanced Computing, Zhengzhou Science and Technology Institute, Zhengzhou, China; email: xiangyangluo@126.com
Ting Wu and Ming Xu are with School of Cyberspace, Hangzhou Dianzi University, Hangzhou, China.
Zhenxing Qian is with the Shanghai Institute of Intelligent Electronics & Systems, School of Computer Science, Fudan University, Shanghai, China.

1. In this paper, DCT coefficients refer to as quantized ones.

It should be noted that the discussion of this paper about steganography/steganalysis is based on JPEG images due to its prevalence. Besides, we establish a statistical model generally under the assumption that the secret bits are embedded in DCT domain [2]. The contributions of this paper are as follows:

- We develop a general framework of statistical model-based detectors, and meanwhile address both strength and weakness of the proposed framework. Under the proposed general framework, one can design a statistical model-based detector, specially aiming at detecting a stego image generated by a typical steganographic scheme.
- This paper empirically analyzes the importance of three different independent factors, referring to as the statistical model, the distribution parameters and payload estimation. Also, the linkage between detection performance and our proposed framework including three factors are illustrated.
- The DCT channels of JPEG images are analyzed in details, in which the fitting performance of each channel is evaluated. Subsequently, in virtue of the strategy of weight assigning, two types of adaptive statistical model-based detectors are devised, respectively with known and unknown channel.

The remainder of this paper is organized as follows. Section 2 mainly overviews the current arts. Section 3 proposes a channel-selected steganographic algorithm to challenge the reliability of the current model-based detectors. The

2. For simplicity, we adopt the term *statistical model* in replace of the term *statistical model of quantized DCT coefficients*.

1545-5971 (c) 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TDSC.2019.2962672, IEEE Transactions on Dependable and Secure Computing

general framework of the statistical model-based detectors is formulated in Section 4. Then, Section 5 mainly investigates the detection performance of current model-based detectors under the general framework, and analyzes the fundamentality of different factors imposing on the detection power. To improve the reliability of statistical detectors, we establish two types of adaptive statistical model-based detectors. Finally, Section 7 presents experimental results of the proposed steganalyzers on the real images, and Section 8 concludes this paper.

2 Related Work

In the face of the challenge proposed by steganography, many steganalytic detectors have been proposed, which can be arbitrarily formulated into two categories: *universal* and *special* ones. The specific overview is extended as follows:

1) The methodologies in the classification of universal detectors mainly rely on extracting features from an inspected image and a trained steganalysis classifier. Generally, the labeled images in the training stage are used for training a detector. Authors of [1], [2], [3] developed a series of universal steganalyzers based on supervised learning mechanism, among which the SVM or ensemble classifier has been the most widely-adopted, and meanwhile deep learning-based classifiers have been designed (see [4], [5]). The study of the universal supervised detectors always advances with the development of steganography. For instance, the current trend of steganography shifts to investigate *adaptive* embedding algorithms, instead of randomly selecting the ready-to-embedded pixels. Due to the less detectable property of the textured or edge zone from the cover image, the adaptive steganographic algorithms have been devised, such as HUGO (Highly Undetectable steGO) [6], [7], WOW (Wavelet Obtained Weights) [8], or UNIWARD (UNIversal WAvelet Relative Distortion) [9]. Meanwhile, universal detectors have been established by extracting universal steganalytic features, such as SPAM (Subtractive Pixels Adjacency Model) [10], SRM (Spatial Rich Model) [11], GFR (Gabor Filter Residuals) [12], and the selection of rich model features [13]. In recent works (see [14], [15]), by assigning different weights based on the embedding possibilities of pixels, the well-designed steganalysis features have been proposed to establish adaptive detectors.

   Based on high dimensional features and supervised learning mechanism, current universal detectors indeed achieve high detection accuracy, that is remarkably better than most of unsupervised steganalyzers. However, the performance of universal steganalysis heavily relies on the features, which should be sensitive to changes caused by steganography. Besides, if the training data is not perfect representative of the cover model, the accuracy of steganalysis cannot be guaranteed, possibly leading to the cover-source mismatch problem [16]. Thus, the problem of designing suitable features for universal steganalysis remains open. Besides, notice that any universal detector cannot be an omnipotent one, that is capable of analyzing all possible steganographic algorithms.

2) The methods in the classification of special detectors aim at detecting one type of known steganography. The weighed stego-image (WS) detector [17] and the test proposed in [18] have verified the effectiveness of detecting LSB replacement steganography. By modifying pixel predictors, adjusting weighting factors and considering bias correction, the improved WS detector was proposed (see [19]). Moving the replacement steganalysis from the spatial to frequency domain, the steganalyzer for JPEG covers proposed in [20] was designed for detecting JSteg steganography. However, it does not allow to get high detection power for a low False Positive Rate (FPR) (see [21]), and its statistical properties remain unknown, which prevents the guarantee of a prescribed FPR. The detector of [22] was an interesting alternative; however it is based on the assumption that DCT coefficients are independent and identically distributed (i.i.d.) within a sub-band, and have a zero expectation which might be inaccurate. By challenging the assumption of i.i.d. coefficients, the improved method (see [23], [24]) designed a statistical detector based on *residual noise* extracted from DCT coefficients. [25] addressed the effect of different denoising filters, optimized the prior statistical detectors, and enhanced the detection performance. Although the detection accuracy was enhanced, the distribution of residual noise can be affected by image content. Therefore, the selection of the optimal denoising filter, which directly determines the accuracy of the distribution model and uncovers the changes of the residual noise caused by embedding, remains open. Besides, the utilization of the accurate joint-distributed or mix model describing DCT coefficients is an alternative solution for detection, which was verified in [26].

   Compared with universal steganalysis, statistical model-based detectors, one type of special detectors, have the particular advantages: 1) the statistical model can help steganalytic investigators analytically study the changes arisen by secret information embedding, instead of empirical analysis of which supervised detectors mainly dependent; 2) cast into the framework of hypothesis testing theory, the performance of the established Likelihood Ratio Test (LRT) can be theoretically analyzed, resulting in that we can successfully achieve the theoretical upper bound of detection at the prescribed FPR; 3) With the estimation of model parameters, the designed practical statistical model-based detector can test an inquiry image without utilization of any prior-trained mechanism like supervised steganalysis.

1545-5971 (c) 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TDSC.2019.2962672, IEEE Transactions on Dependable and Secure Computing

Even though the study of the statistical model-based detectors has been investigated for many years, the limitation of the framework, and the relationship between the detection result and the detector parameters remain unknown. More importantly, to deal with the problem of adaptive steganography, the design of adaptive statistical model-based detector (not relying on a large scale of training set) with improved performance is a challenging task. Thus, in this paper, we propose to design adaptive statistical model-based detectors for dealing with those problems. For clarity, it is proposed to summarize the main notations used in this paper in Table 1.

TABLE 1: Notations

$\mathbf{I}$	Inspected image
$\mathbf{C}$	Unaltered cover image
$\mathbf{S}$	Stego image containing a secret message
R	Payload
$\mathcal{H}_0$ and $\mathcal{H}_0$	Two hypotheses
$\mathcal{P}_{\boldsymbol{\theta}}$	Distribution followed by cover source
$\mathcal{Q}_{\boldsymbol{\theta}}^{R}$	Distribution followed by stego source
$\mathcal{K}$	Class of tests
$\boldsymbol{\theta}$	Model parameters
α_0	False Alarm Rate (FAR)
β_δ	True Positive Rate (TPR)
Λ	Likelihood Ratio (LR)
δ	Statistical model-based detector
$\widehat{\delta}$	Adaptive statistical model-based detector
ω_l	Weight assigned for channel l

3 Limitation of Current Model-Based Steganalysis

In the practical design of steganography, the undetectability of steganography is usually given the priority. Thus, the steganographic schemes should be constructed in the guidance of the following principals (see Chapter 7 of [27] for details):

1) The model of the cover source is perfectly preserved after embedding;
2) The procedure of embedding behaves as (or very similarly to) natural image process;
3) The steganographic algorithm is capable of resisting known steganalysis attacks;
4) The impact of embedding is minimized.

In modern steganography, to minimize the distortion caused by embedding (Principle 4), some adaptive (or channel-selected) steganographic schemes have been adopted, and capable of degrading the detection power of modern detectors (see [6], [7], [8], [9] for instance). Inspired by the strategy of channel-selection, together with the guidance of principals proposed by [27], we intend to design a steganographic scheme capable of resisting current statistical model-based steganalysis attacks (Principal 3). Specifically, statistical model-based detectors are established mainly dependent of the assumed distribution model. Once the accuracy of the model (used for describing the features extracted from an inquiry image) cannot be guaranteed, the performance of model-based detectors are also probably degraded.

In this section, by selecting DCT channels, which are difficult to statistically model, we propose a novel channel-selected algorithm with the ability of challenging a statistical model-based detector. Furthermore, the effectiveness of the proposed channel-selected steganography is verified through empirical results. It is worth noticing that the aim of this paper is not only to design a channel-selected steganographic algorithm for exposing limitation of current model-based steganalysis, but also to establish more reliable adaptive statistical model-based detectors with high detection performance (see details in Section 6). In fact, the techniques of both steganography and steganalysis are mutually advanced.

3.1 Design of Channel-selected Steganography

Inspired by the non-channel-selected JSteg steganography (see [28]), to explore the limitations of current statistical model-based detectors, let us design the embedding algorithm by selecting DCT channels. The principle of the algorithm is to replace the LSB of DCT coefficients by bits of the message to be hidden. In the JPEG compression scheme, the DCT transform is applied to blocks of 8×8 pixels. Then for each channel $l \in \{1, \ldots, 63\}$ except channel 0, we propose to embed the bit stream into the selected channels, referring to as the middle or high frequency channels which are difficult to statistically model. If the embedding bit (0 or 1) matches the LSB of DCT coefficient, no change is made; if not, the LSB of DCT coefficient is flipped to 1 or 0. Besides, it should be noted that the channels are selected based on the fitting performance (see details in Section 6.1).

It is reasonable that the selected channel is distributed in the middle or high frequency, where the selected DCT coefficients can be mapped to high-textured regions in the spatial domain. In fact, considering the undetectability of embedding secret messages into a cover image, the steganographer probably embeds the secret bits into the high-textured or edge regions with lower detectability than smooth regions.

3.2 Empirical Results

The benchmark BOSSbase dataset [29] including 10000 uncompressed grey-level images with 512 $\times$ 512 pixels is used for our experiments. It is proposed to adopt minimal Probability of Error (P_E) as metric, which corresponds to the minimal value of false positive and false negative rate, and is formally defined as:

$$P_E = \min_{\alpha_0 \in (0,1)} \frac{\alpha_0 + (1 - \beta_\delta)}{2}. \tag{1}$$

where α_0 denotes the FPR while $1 - \beta_\delta$ represents the false negative rate (see detailed description in Eqs. (6) (7)). Thus, the larger value of P_E implies the worse performance of the steganalysis detector.

For clear comparison, when the payload (see Eq. (4)) and quality factor (QF) are both prescribed, two P_E's are illustrated in Table 2, in which the first column represents

1545-5971 (c) 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TDSC.2019.2962672, IEEE Transactions on Dependable and Secure Computing

TABLE 2: Minimal P_E comparison, using the proposed channel-selected (CS) or non-channel-selected (Non-CS) steganographic algorithm (corresponding to the results of the second column with *italic* typeface); the statistical model-based detector using the algorithm of [22].

Payload \ QF	85	90	95
0.05	0.43 *0.38*	0.45 *0.40*	0.48 *0.44*
0.10	0.37 *0.29*	0.41 *0.33*	0.45 *0.39*
0.20	0.28 *0.17*	0.34 *0.23*	0.41 *0.32*
0.30	0.22 *0.10*	0.28 *0.16*	0.36 *0.26*
0.40	0.17 *0.06*	0.23 *0.11*	0.33 *0.21*
0.50	0.14 *0.04*	0.20 *0.08*	0.29 *0.17*
Average	0.27 *0.17*	0.32 *0.22*	0.39 *0.30*

the channel-selected steganography, and the second column (*italic* typeface) for non-channel-selected steganography. From Table 2, it can be observed that at the given payload 0.05, the P_E approaches 0.5 for JPEG images with three quality factors. When the P_E equals to 0.5, the statistical detector [22] is invalid as random guess. In fact, besides [22], when suffering the challenge from CS steganographic algorithm, other reliable statistical detectors [25], [26] and even the supervised powerful detector [3] also have worse performance (see Tables 4, 5, and 8 for details)

It is observed that with increasing the payload, the performance of the detector is gradually improved. Because the more secret bits are embedded into the cover image. In addition, if the larger quality factor is utilized for compression, the undetectability of the stego image can be further improved. In fact, with increasing the quality factor, the redundancy of the compressed image is increased. Meanwhile, the high correlation among pixels of an inquiry image unavoidably disturb the differences caused by embedding. Therefore, the performance of the statistical model-based detector cannot be guaranteed. It should be noted that for some supervised universal detectors, for instance the JSRM-based detector (see [9], [30] for details), the similar property of detection performance was also exposed.

When the channel-selected strategy is used, it is difficult to detect the stego image using the current model-based detectors (see extensive results in Section 7.2). Hence, for detecting the channel-selected steganography, it is imperative to re-investigate the design of statistical model-based detectors, and devise the more reliable adaptive detector.

In this paper, we first formulate the general framework of current statistical model-based detectors, that is empirically analyzed with different independent *factors*, referring to as the statistical model, the distribution parameters and payload estimation. Then, more importantly, a novel adaptive statistical model-based detector is designed that is more reliable in the face of both channel-selected and non-channel-selected steganography.

4 General Framework of Statistical Model-based Detector

Recently, a series of statistical model-based detectors have been established, among which some are designed for LSB replacement (see [22], [23], [24], [25], [26], [31]); the others are for LSB matching [32]. In the face of challenges proposed by various steganographic algorithms, a statistical model-based detector can only be applied for one typical steganography while not for universal detection. Although many models has been designed to generate statistical steganalysis detectors in the prior studies, few of them are unified into the general framework. In this context, we propose to establish the general framework, which can guide us to design a steganalysis detector based on a statistical model. More importantly, under the proposed general framework (see Section 4), we can comprehensively investigate the relationship between detection performance and our proposed framework containing three factors (see Section 5), further resulting in the establishment of the adaptive detectors (see Section 6).

For simplicity, let us first formulate the general framework of designing a statistical model-based detector. To our knowledge, current statistical model-based detectors can be formulated under the framework of hypothesis testing theory. Therefore, when analyzing an inspected image $\mathbf{I} = \{i_n\}$, $n \in \{1, \cdots, N\}$, a steganalytic researcher usually decides between the following two hypotheses:

$$\begin{cases} \mathcal{H}_0 : \mathbf{I} = \mathbf{C} & \text{is an unaltered cover image} \\ \mathcal{H}_1 : \mathbf{I} = \mathbf{S} & \text{is a stego image containing a secret message,} \end{cases} \tag{2}$$

where an unaltered cover image denoted as $\mathbf{C} = \{c_n\}, n \in \{1, \ldots, N\}$, in which N denotes the total number of pixels; a stego image as $\mathbf{S} = \{s_n\}, n \in \{1, \ldots, N\}$, where again N denotes the total number of pixels. To deal with the problem of binary classification, a statistical model-based detector distinguishes the image under investigation in virtue of its statistical distribution which the pixels (in the spatial domain) or coefficients (in the frequency domain, such as DCT domain for JPEG images) follow. Therefore, let us transfer the study of detector to the research of the statistical distribution of image pixels/coefficients. For simplicity, in the following study, we utilize "pixels" to denote both "pixels/coefficients".

Let us assume that $\mathbf{I} = \{i_n\}$ are independent, and they all follow the probability distribution, denoted as $\mathcal{P}_{\boldsymbol{\theta}}$, parameterized by the parameter $\boldsymbol{\theta}$. By considering that the pixels are quantized, the distribution $\mathcal{P}_{\boldsymbol{\theta}}$ is represented by its probability mass function (pmf) denoted $P_{\boldsymbol{\theta}} = \{p_{\boldsymbol{\theta}}[i_n]\}$. Suppose that a cover image $\mathbf{C}$ is used as a carrier to hide a secret message with payload R. The pixels from the stego image, denoted $\mathbf{S}$, follow the statistical distribution $\mathcal{Q}^R_{\boldsymbol{\theta}}$ which is completely characterized by its pmf $Q^R_{\boldsymbol{\theta}} = \{q^R_{\boldsymbol{\theta}}[i_n]\}$ given by:

$$q^R_{\boldsymbol{\theta}}[i_n] = (1 - \frac{R}{2}) p_{\boldsymbol{\theta}}[i_n] + \frac{R}{2} p_{\boldsymbol{\theta}}[\overline{i_n}], \tag{3}$$

where, again, $\overline{i_n}$ represents the integer i_n with flipped Least Significant Bit (LSB). Since the secret message bits, which are i.i.d, follow a Binomial distribution $\mathcal{B}(1, 1/2)$, the probability that a stego pixel s_n equals the cover c_n with flipped LSB equals $R/2$ while, on the contrary, the probability of a stego pixel s_n equal to the cover pixel c_n equals $1 - R/2$. It should

1545-5971 (c) 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TDSC.2019.2962672, IEEE Transactions on Dependable and Secure Computing

be noted that the payload R represents relative embedding rate. It can be defined as:

$$R = \frac{L}{\sum i_u} \tag{4}$$

where $\sum i_u$ denotes the number of usable (can be used for embedding the secret message) pixels; L represents the length of secret message bits. For instance, if we adopt a JSteg embedding algorithm, the accounted number of i_u has to exclude all the DC coefficients or AC coefficients equal to "0" and "1". In practice, for each cover source, it probably has different embedding capacity. Because different cover sources might have different useable pixels. Therefore, it makes sense that we use the relative payload R (see Eq. (4)) to evaluate the performance of steganography and steganalysis.

To specify the framework of binary classification (2), a statistical steganalyst is prone to make a choice between the following two hypotheses: $\mathcal{H}_0$: "*the pixels i_n follow the distribution $\mathcal{P}_{\boldsymbol{\theta}}$*" and $\mathcal{H}_1$: "*the pixels i_n follow the distribution $\mathcal{Q}^R_{\boldsymbol{\theta}}$*" which can be written formally as:

$$\begin{cases} \mathcal{H}_0 : \{\mathbf{I} \sim \mathcal{P}_{\boldsymbol{\theta}}\}, \\ \mathcal{H}_1 : \left\{\mathbf{I} \sim \mathcal{Q}^R_{\boldsymbol{\theta}}\right\}. \end{cases} \tag{5}$$

A statistical test is a mapping $\delta : \mathbb{I}^N \mapsto \{\mathcal{H}_0, \mathcal{H}_1\}$, and focuses on the Neyman-Pearson bi-criteria approach (see [33] for details): maximizing the correct detection probability for a given FPR α_0. Let:

$$\mathcal{K}_{\alpha_0} = \left\{ \delta : \sup_{\boldsymbol{\theta}} \mathbb{P}_{\mathcal{H}_0}[\delta(\mathbf{I}) = \mathcal{H}_1] \leq \alpha_0 \right\}, \tag{6}$$

be the class of tests with a FPR upper-bounded by α_0. Here $\mathbb{P}_{\mathcal{H}_i}[A]$ represents the probability of event A under hypothesis $\mathcal{H}_i, i = \{0, 1\}$, and the supremum over $\boldsymbol{\theta}$ has to be understood as whatever the distribution parameters might be, guaranteeing that the FPR α_0 can not be exceeded. Among all the tests in $\mathcal{K}_{\alpha_0}$, it is aimed at finding a test δ with maximizing the power function, defined by the True Positive Rate (TPR):

$$\beta_\delta = \mathbb{P}_{\mathcal{H}_1}[\delta(\mathbf{I}) = \mathcal{H}_1], \tag{7}$$

which is equivalent to minimize the false negative rate $\alpha_1(\delta) = \mathbb{P}_{\mathcal{H}_1}[\delta(\mathbf{I}) = \mathcal{H}_0] = 1 - \beta_\delta$.

To solve the problem of finding the optimal test $\delta(\mathbf{I})$, a statistical detector is generally designed relying on the LRT, which can be expressed by:

$$\delta(\mathbf{I}) = \begin{cases} \mathcal{H}_0 \text{ if } \Lambda(\mathbf{I}) = \dfrac{Q^R_{\boldsymbol{\theta}}[\mathbf{I}]}{P_{\boldsymbol{\theta}}[\mathbf{I}]} < \tau \\ \mathcal{H}_1 \text{ if } \Lambda(\mathbf{I}) = \dfrac{Q^R_{\boldsymbol{\theta}}[\mathbf{I}]}{P_{\boldsymbol{\theta}}[\mathbf{I}]} \geq \tau, \end{cases} \tag{8}$$

where Λ denotes the Likelihood Ratio (LR) and τ a decision threshold set to guarantee a prescribed FPR. To detect the stego image with secret bits hidden in the spatial domain, each observation (or pixel) i_n of $\mathbf{I}$ follows the distribution, denoting $\mathcal{Q}^R_{\boldsymbol{\theta}}$ associated with its pmf $Q^R_{\boldsymbol{\theta}}$ while cover source is modelled by $\mathcal{P}_{\boldsymbol{\theta}}$ with its pmf $P_{\boldsymbol{\theta}}$. Similarly, in the case that data hiding happens in DCT domain for JPEG images, each observation, referring to as a DCT coefficient, is modelled by $\mathcal{Q}^R_{\boldsymbol{\theta}}$ under hypothesis $\mathcal{H}_1$ or $\mathcal{P}_{\boldsymbol{\theta}}$ under hypothesis $\mathcal{H}_0$.

Generally, under the general framework of designing a statistical model-based detector, we first need to formulate the LRT. In that case, it is assumed that all the model parameters can be obtained before testing. The reason why we study the LRT is that we need to testify the effectiveness of the proposed distribution model, involving both $\mathcal{Q}^R_{\boldsymbol{\theta}}$ and $\mathcal{P}_{\boldsymbol{\theta}}$. Since pixels/coefficients extracted from digital images are generally heterogenous, among each authentic image (cover or stego), we cannot guarantee that each pixel can be modelled without any error. With the help of Montecarlo simulation, the random variables, simulating the real pixel values, follow the assumed distribution (see [25] for instance). By excluding the interference from prediction errors of parameter $\boldsymbol{\theta}$ plus mismatch existing between empirical distribution of random variables and the assumed statistical model, the LRT allows to achieve the optimal detection power. Moreover, the normalized LRT can establish the theoretical performance of the proposed statistical detector. At the given FPR, the LRT is capable of generating a theoretical upper bound of detection. To study the specific discussion of the LRT, readers may refer to [25].

Following the design of the LRT, under the framework of establishing a statistical model-based detector, it is proposed to investigate the real practical detector for steganalysis. To deal with the problem of transferring the LRT to the practical (or real) detection, we have to solve one problem: estimation of parameter $\boldsymbol{\theta}$. If parameters are estimated based on the MLE algorithm (see [22]), the GLRT can be directly established; if parameters are obtained with the help of other algorithms (see [23]), such as convolution property of denoising filters, the Practical Likelihood Ratio Test (PLRT) can be obtained. In fact, whatever the GLRT or PLRT is utilized by steganalytic researchers, the real practical detector is indeed evolved from its corresponding LRT by considering the problem of parameters estimation.

In addition, in the practical detection, another parameter payload R cannot be acquired before steganalytic testing. It plausibly makes sense that the errors of payload estimation can result in the perturbation of detection accuracy of a statistical model-based detector (see [16]). However, in this context, it is proposed to challenge that assumption. The specific analysis is extended in Section 5.3.

Currently, a series of statistical model-based detectors has been continuously established [22], [23], [24], [25], [26], [31], [32], [34], dealing with some steganographic algorithms such as LSB replacement or matching. In the following section, it is proposed to primarily analyze the performance of statistical model-based detectors, which detect the algorithm of data hiding happening in DCT domain of JPEG images for instance. In fact, our analytic results can be extended to the spatial or other frequential domain.

5 Performance of Statistical Model-based Detector under General Framework

The development of steganalysis always follows the advancement of steganography. In recent study of steganography, steganographers are prone to design the embedding algorithm by minimizing the universal distortion function of an inspected image under the framework of Syndrome trellis Coding (STC) [9], which can be arbitrarily defined as

1545-5971 (c) 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TDSC.2019.2962672, IEEE Transactions on Dependable and Secure Computing

adaptive steganography. With the embedding payload approaching the theoretical upper bound, a series of steganographic algorithms applied in the spatial or frequential domain is proposed to improve the efficiency of the algorithm [9]. However, most of typical steganographic tools spread over the Internet are still designed based on the traditional techniques by considering practicability (easy to realize) and efficiency. Thus, in this practical context, it is proposed to utilize the LSB replacement methodology as the embedding paradigm, for instance in DCT domain of JPEG images. Then, we mainly analyze the detection performance of the statistical model-based detector.

Under the general framework proposed in Section 4, the general steps of designing an efficient statistical model-based detector are as follows: 1) selecting the optimal model describing the distribution of pixels (in the spatial domain) or coefficients (in the frequential domain, such as DCT domain); 2) estimating model parameters; 3) predicting payload. In the following subsections, let us investigate the above three factors; meanwhile we demonstrate the relationship between factors and the detection power of the statistical model-based detector.

5.1 Statistical Model

In practice, several distributions have been proposed in the literature to model the DCT coefficients. Some accurate models such as the Generalized Gaussian [35] and, more recently, the Generalized Gamma model [36] have been shown to provide more accurate description of DCT coefficients' distribution at the cost of higher complexity. Some of those models have been exploited in the field of steganalysis, see [37], [38] for instance. However, to the best of our knowledge, the Laplacian distribution is probably the most widely-used due to its simplicity and fairly good accuracy.

Immediately, it is proposed to give the probability density function (pdf) of the continuous Laplacian distribution:

$$f_{\boldsymbol{\theta}}(x) = \frac{1}{2b} \exp\left(\frac{-\mid x-\mu \mid}{b}\right), \forall x \in \mathbb{R} \tag{9}$$

where the model parameter $\boldsymbol{\theta}$ involves location parameter μ representing the expectation of the Laplacian distribution, and b the scale parameter denoting the variance. Under the assumption of the Laplacian distribution, $\mathcal{P}_{\boldsymbol{\theta}}$ is the discrete description of Eq. (9) while $\mathcal{Q}_{\boldsymbol{\theta}}^{R}$ can be straightforwardly obtained (see Eq. (3)). Then it is required to represent the random variable x using the pixel of the real inspected image.

Without the loss of generality, let us denote a grey-level image with the size $M \times N$ as $\mathbf{I}(m,n)$ the pixels intensity of a given image with (m,n) the pixel position, where M and N denote the height and the width of the image $\mathbf{I}(m,n)$. In this context, we mainly investigate the detector of detecting an inquiry image with secret bits embedded in DCT domain. Thus, Let us define the DCT coefficients by the matrix $\mathbf{D} = \{d_{m,n}\}$. The DCT coefficients of the image $\mathbf{I}$ is acquired usually over 8×8 blocks. Since, in the procedure of designing a statistical detector, the model is generally established in virtue of the same DCT sub-band, referring to the same position of each DCT 8×8 block, let us re-define the DCT coefficients by the matrix $\mathbf{V}_{k,l}$, $k \in \{1, \ldots, K\}$, $l \in \{0, \ldots, 63\}$ with $K \approx M \times N/64$. In this context, we assume that both width and height of an inspected image are multiples of 8. In the assumed distribution, the random variable x of Eq. (9) can be replaced by each DCT coefficient $v_{k,l}$ among the same sub-band l.

In the framework of optimal detection (see Eq. (8)), the first attempt has been made to design a statistical test modelling the DCT coefficient with the Laplacian distribution (see [22]). It was assumed that the Laplacian model was optimal, taking into account the computation cost and the complexity of the distribution. However, by challenging the conventional Laplacian model, [26] utilized the mix model (or improved version of the traditional Laplacian model) proposed in [39] for establishing steganalysis detector. After selecting the optimal model, estimation algorithm is studied; denoising filters were utilized for parameters estimation, assuming that coefficients are not i.i.d (see [23], [24], [25]). Last, under the framework of hypothesis testing theory, the statistical detector can be established based on the value of Likelihood Ratio (LR). Because the LR-based test can achieve the optimal detection power at the given FPR.

The designed statistical model-based detector has verifies that it can achieve the satisfying detection performance at the given FPR. However, the practical TPR acquired by using the statistical detectors still cannot match the theoretical results. Because one cannot guarantee that the proposed model perfectly describes the distribution of all DCT coefficients from each inquiry image, that leads to the prediction error during parameter estimation. In real steganalysis, a large scale image dataset contains various samples with heterogenous property of pixels, that indeed raises the difficulty of modelling pixels or coefficients of each inquiry image using the statistical model.

In general, the prediction error can be attributed as follows: the DCT coefficients are not independent of image content, which hardly follow the Laplacian distribution when an inquiry image contains large smooth regions; a mass of zero-value DCT coefficients impact the accuracy of parameter estimation, especially in the case of the high frequency domain. That limitation of the statistical model-based detector directly inspires us to design the channel-selected steganographic algorithm (see Section 3.1 for details).

In addition, the statistical model is the significant footstone (or deterministic factor) of the statistical model-based detector; the accuracy of the estimated parameters determines the robustness of the detector; the payload estimator to some extent controls degree of accuracy. In the following, it is proposed to specifically analyze the parameters of the statistical model. Then, let us continue extending our discussion about the impact of payload estimation on the detection power.

5.2 Effects of Model Parameters Selection

Let us directly illustrate the specific experimental results, that are more solid than the sophisticated and boring description. First, all the uncompressed images from the benchmark BOSSbase dataset are compressed in JPEG format as cover carriers with quality factor 70. Then, 10000

1545-5971 (c) 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TDSC.2019.2962672, IEEE Transactions on Dependable and Secure Computing

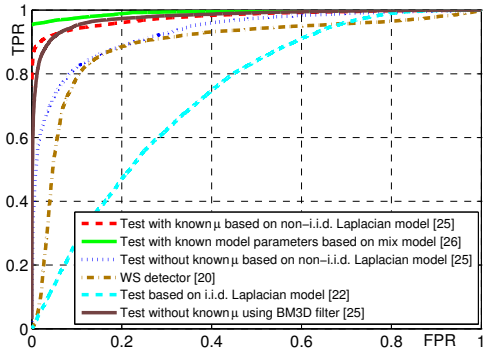


Fig. 1: ROC comparison of detection performance from BOSSbase dataset with quality factor 70.

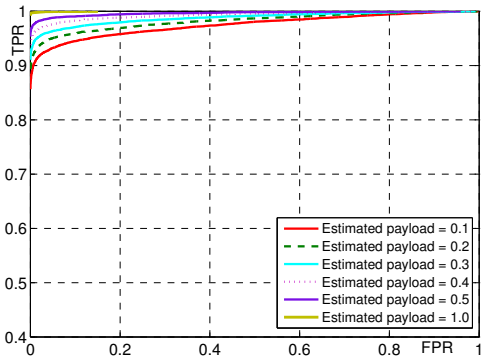


Fig. 2: ROC comparison of detection performance from BOSSbase dataset with quality factor 85; all the stego images are embedded with authentic payload 0.3.

stego images with payload 0.05 are obtained using a typical JPEG steganographic scheme. It should be noted that regardless of data hiding algorithms, we only discuss the relationship between detector performance and the statistical model, and its parameters.

In the experiments, let us compare some typical detectors of detecting JPEG steganography. Mainly relying on our experimental results, the accuracy of the statistical detectors is dominated by two significant factors: distribution model and estimated parameters. For simplicity, all the detectors are analyzed for detecting JSteg steganography. Again, it is worth noting that the analysis results can guide us to establish more effective and reliable statistical model-based detectors.

As Fig. 1 illustrates, the mix model-based detector [26] achieves the best result. Because the mix model [39] describes the distribution of DCT coefficients better than other models. Since it hardly holds true that all the DCT coefficients in one sub-band follow the Laplacian model with the same parameters, the test based on i.i.d. Laplacian model [22] performs worst. Besides, the WS detector [20] assuming that DCT coefficients asymptotically follow the Gaussian distribution also does not perform very well.

The results from the test with/without known μ intrigue us most. Supposing that our model has been confirmed, the problem of how to improve detection accuracy of the designed test is very attractive. In this case, we assume that each coefficient has its unique parameter, not sharing the same parameter among the sub-band. Therefore, it is proposed to estimate the corresponding parameter of each coefficient. One may argue that although the non-i.i.d. model probably helps us improve the accuracy of the model for describing the DCT distribution, it might unavoidably introduce the accumulated prediction errors. Conversely, as Fig. 1 demonstrates, compared with the i.i.d. Laplacian model [22], the non-i.i.d. model-based detector [25] can obtain considerably better ROC curve. That result empirically verifies that the better-performed detector with limited prediction errors can be acceptable. Besides, it should be noted that authors of [25] only proposed the non-i.i.d. model, but did not empirically analyze the model parameter μ, which this context focuses on.

Besides, in the assumption that the Laplacian model is non-i.i.d, when the detector has known parameter μ, the detector is considerably improved. In fact, Fig. 1 verifies that before and after embedding secret bits, if parameter μ (denoting the expectation value of the inspected image) remains the same, the detection power can be largely enhanced. In this scenario, the secret bits, referring to as additive noise with uniform distribution, nearly cannot modify the expectation value of the image. Thus, the problem of estimating parameter μ is with respect to the design of an ideal denoising filter. Then, the ideal denoising filter need to guarantee that the estimated parameter μ remains the same before and after embedding. In our prior research (see [25]), the BM3D denoising filter is the current optimal choice. In fact, we believe that other filters (better than BM3D denoising filter) indeed exist, whose performance might nearly equal to the result of the ideal one. Nevertheless, in this context, we provide the illustrative exposition of optimizing a denoising filter used for the design of a statistical model-based detector.

The test with BM3D adopts the non-i.i.d. Laplacian model. Different from the test without known μ (using wavelet denoising filter for estimating parameter μ), the test with BM3D utilizes BM3D denoising filter for estimation. The ROC curve implies that a more accurate parameter directly results into better detection performance. However, even we select BM3D denoising filter, its performance (at the very small FPR) is a little worse than the test with known μ. This result verifies that secret bits as additive noise unavoidably disturb parameter estimation. Therefore, we conclude that by selecting the optimal denoising filter for estimating parameter, we can furthermore improve the detection power to some extent, yet not exclude the impact of embedding.

5.3 Discussion of Payload Estimation

In this subsection, we mainly discuss the impact of payload estimation on detection accuracy. [16] exposed the disadvantages of the statistical model-based detector, referring to as unknown payload. The algorithm of payload estimation has been studied, such as [40] and [26]. Besides, the Locally Asymptotically Uniformly Most Powerful (LAUMP) is an alternative solution, that has been investigated in [32]. When the steganographic tool can be acquired, the estimator proposed in [14] can be designed by randomly re-embedding the same message for multiple times. In the design of the

1545-5971 (c) 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TDSC.2019.2962672, IEEE Transactions on Dependable and Secure Computing

statistical detector, the payload serves as one unknown parameter of the detector (see Eqs. (3) and (8)). Thus it is plausible that the accuracy of the estimated payload is obligatory. However, in this context, it is proposed to challenge that assumption.

It is proposed to directly testify our assumption using the solid experimental results. Let us utilize 10000 JPEG images as cover objects with quality factor 85 of the BOSSbase dataset, which is used to generate JSteg-based stego images with payload 0.3. Then, in the procedure of steganalysis, the estimated payload of the statistical detector is expanded from 0.1 to 0.5. As Fig. 2 illustrates, the estimation of payload hardly interferes with the performance of the statistical model-based detector. The estimated payload ranging from 0.1 to 0.5 neighboring around the authentic payload 0.3 indeed helps the statistical model-based detector achieve relevant results. Thus, for a statistical detector, the payload estimation is not a deterministic factor. It should be noted that in this experiment, we adopt the detector based on non-i.i.d. model with known parameter μ. In the practical steganalysis, the statistical model-based detectors are not very easy to establish, considering the distribution model and parameter estimation which the steganalysis researchers should mainly put focus on. Nevertheless, the study of payload estimation is not very related to the fundamentality of the statistical model-based detector.

In virtue of the probability $Q_{\boldsymbol{\theta}}^{R}$ (see Eq. (3)), let us then analyze the detection performance of the detector the payload acts on. In real steganography, to avoid the detection, the steganography designer is always trying to prescribe a very small payload, about 0.05 for instance. In that case, for the probability of each random variable, the $q_{\boldsymbol{\theta}}^{R}$, R actually does not donate much to the overall probability. Thus, the tiny mismatch between the authentic and estimated payload can be to some extent tolerated. Moreover, even if the tiny mismatch happens, we are prone to use the larger value to offset the prediction error, that is empirically verified in Fig. 2.

Based on our discussion in Section 5, in the procedure of establishing the statistical model-based detector under the general framework, while designing the well-performed detector with high accuracy, the statistical detector requires: 1) the nearly-perfect model; 2) the accurately estimated parameters (or selection of an optimal filter) within the limitation of prediction errors; 3) payload estimation within tolerable errors.

6 Adaptive Statistical Model-based Detector Designed for JPEG Images

Although the training-free statistical model-based detectors have been studied while it cannot deal with the case of adaptive steganography (see Table 2 of Section 3 for instance). To improve the reliability of the statistical model-based detectors and to counter the proposed channel-selected steganography, we intend to devise adaptive statistical model-based detectors. First of all, it is proposed to analyze the DCT channels of JPEG images. Then based on the analysis, relying on our proposed general framework combined with the mechanism of weight assigning, we establish two types of adaptive detectors, respectively with known and unknown DCT channel.

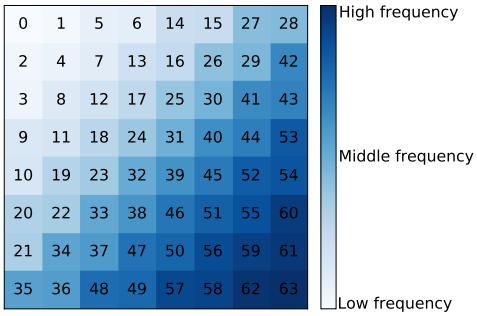


Fig. 3: DCT channel index using Zig-zag order with gradually increasing the number, corresponding to different frequency intensity, from low frequency to high frequency.

6.1 Investigation of DCT Channels

Since each inquiry image might have different texture characteristics, resulting into that not all DCT channels of the image can perfectly follow the proposed model, the remarkable limitation of the model-based detectors is exposed. Then, we need to study the fitting results of the proposed statistical model for each channel.

Let us conduct the χ^2 Goodness-of-Fit (GOF) measuring the model fitting results. Here, the χ^2 statistic values is defined by:

$$\chi^2 = \sum_{i=1}^{N} \frac{(e_i - t_i)^2}{t_i} \tag{10}$$

where i denotes the bin index, ranging from index 1 to 63 excluding index 0 following Zig-zag order (see Fig. 3), of the DCT histogram on the each channel, N the number of bins, e_i defined as the counts of each bin, referring to as the empirical value, t_i the expectation value calculated using the pdf of the assumed model. In our analysis, it is proposed to randomly select 200 images with the size 512×512 from the BOSSbase dataset, that are compressed with QF 70. The averaged χ^2 can be obtained. It should be noted that each channel consists of 4096 coefficients; the empirical results are normalized.

Fig. 4 reports the average results of each DCT channel over 200 images. The percent illustrated in Fig. 4 represents the ratio of the used number of DCT coefficients in each channel. We observe that the larger χ^2 statistic value is, the more prediction error in each bin between the expected value from the proposed model (the Laplacian model is used) and real empirical value generated from the statistical histogram. Besides, in our analysis, we do not consider the last channels with the highest frequency. Because a large scale of zero is generated from those channels, not used for secret information embedding.

Therefore, in virtue of the experimental results from Fig. 4, let us draw the following two assumptions: 1) with increasing the channel index, the trend of χ^2 value is gradually elevated with some vibration, implying worse and worse performance of the statistical model-based detectors;

1545-5971 (c) 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TDSC.2019.2962672, IEEE Transactions on Dependable and Secure Computing

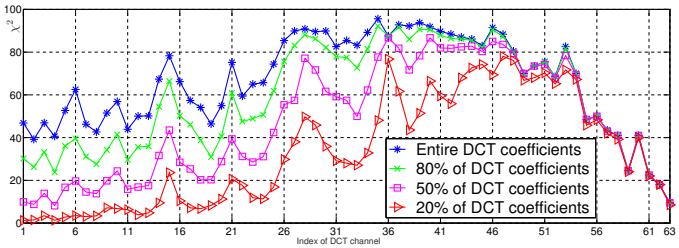


Fig. 4: χ^2 GOF test (averaged from randomly selected 200 images) results following Zig-zag order among each channel with different quantities; the detailed calculation refers to Eq. (10).

2) by using less coefficients (subset of each channel, for instance 80%, 50%, and 20% in Fig. 4), the fitting performance is indeed enhanced. Since in the high-frequency channel with large channel index, many zero-value DCT coefficients appear due to the large quantization step, coefficients cannot be guaranteed to follow the Laplacian model. Besides, if all the coefficients of each channel are used for fitting the proposed model, many outliers of the assumed variables are generated, resulting into poor-performed χ^2 test. Additionally, when we compress the same 200 images with QF 85 or 90, the fitting results are very similar to that of Fig. 4.

Through analyzing the DCT channels, we would like to design the strategy of weight assigning, that is different channels acquire different weights. In this paper, we mainly aim to improve the reliability of current statistical model-based detectors capable of detecting both channel-selected and non-channel-selected steganography. To this end, the specific establishment of adaptive statistical model-based detectors is extended in the following subsections.

6.2 Problem Statement

In the previous design of the statistical detectors, we assume that the DCT coefficients or pixels have the same contribution to establish our detector. To challenge that assumption and to counter the threat from the channel-selected steganography, let us design an adaptive statistical model-based detector. The principle of devising an adaptive detector is that we assign different weights to different DCT channels, instead of the uniform weight. Thus, in virtue of the established general framework (see Eq. (8)), let us establish the adaptive statistical model-based detector formulated by:

$$\widehat{\delta}(\mathbf{V}) = \begin{cases} \mathcal{H}_0 \ \text{ if } \widehat{\Lambda}(\mathbf{V}) = \sum\limits_{k=1}^{K}\sum\limits_{l=1}^{63} \widehat{\Lambda}_{apt}(v_{k,l}) < \widehat{\tau}, \\ \mathcal{H}_1 \ \text{ if } \widehat{\Lambda}(\mathbf{V}) = \sum\limits_{k=1}^{K}\sum\limits_{l=1}^{63} \widehat{\Lambda}_{apt}(v_{k,l}) \geq \widehat{\tau}, \end{cases} \tag{11}$$

where the adaptive statistic $\widehat{\Lambda}_{apt}(v_{k,l}) = \widehat{\Lambda}(v_{k,l}) \cdot \omega_l$ is the realization of the general framework of the statistical model-based detector in DCT domain. In the practical detection, the variable $\widehat{\Lambda}(v_{k,l})$ is the prediction of $\Lambda(v_{k,l})$ with estimated model parameters among the k_{th} coefficient of the channel l. The specific discussion of our proposed framework refers to Section 4.

Without the loss of generality, the design of adaptive detector is strictly under the proposed general framework. By calculating the ratio values based on the assumed statistical model, $\widehat{\Lambda}(\mathbf{V})$ the sum of LR values serving as discrimination factor is used for detecting stego images. However, the prior study such as [26] only considers the DCT coefficients as independent random variables while to some extent ignoring the intrinsic property of them in a stego image. In this scenario, when the adaptive steganographic algorithm embeds hidden bitts in the middle or high frequency where the DCT channels are difficult to model, the detection error is increased (see Fig. 5 for instance). The possible reason is that the discriminability of discrimination factor between stego and cover source become degraded. In other words, in the classification, the maximum margin lying between discrimination factor from different sources is narrowed down, leading to the increasement of detection error. To further enforce the discriminability and improve the detection performance, we propose to refine the selection of DCT coefficients donating much to the discriminability. Specifically, $\widehat{\Lambda}(\mathbf{V})$ consists of each LR value $\widehat{\Lambda}_{apt}(v_{k,l})$ which is sumized as the discrimination factor. In our assumption, if the embedding happens in the position $\{k, l\}$, the LR value of stego source is larger than that of its corresponding cover source. Thus, if the DCT coefficient with high embedding probability is used for embedding, it should be assigned more weights to donate the calculation of LR value, resulting in that the discriminative features are further strengthened. In fact, by accumulating all the LR values with various weights (instead of uniform weight from prior non-adaptive detectors), the discriminability of discrimination factor becomes more effective for steganalysis. Besides, it is worth noting that regardless of statistical models, both the adaptive detector and previous non-adaptive detector such as [26] are unified under our proposed framework. By assigning various weights to DCT channels, the non-adaptive detector is modified as adaptive one; by setting the unit weight for all the channels, the adaptive detector degenerates back to non-adaptive one.

6.3 Adaptive Detector with Known Channel

In fact, the study of designing an adaptive statistical model-based detector can be transferred to formulate the parameter ω_l, which can assign the weight to each DCT channel. We have discussed that the steganographer is prone to embed the secret information into the middle or high frequency channel (see details in Section 3) with considering the undetectability of embedding. Straightforwardly, the adaptive detector should assign the large weights to those middle or high frequency channels. Therefore, let us consider two scenarios. In the first scenario, the steganalyzer clearly knows the used channels for embedding. Immediately, the weight ω_l can be defined as:

$$w_l = \begin{cases} 1 & \text{if} \quad \text{the channel } l \text{ is used for embedding,} \\ 0 & \text{otherwise.} \end{cases} \tag{12}$$

In fact, in this scenario, we only statistically compute the used DCT channels for establishing our statistical detector. In the practical detection, it is possible that the steganalysis

1545-5971 (c) 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TDSC.2019.2962672, IEEE Transactions on Dependable and Secure Computing

can acquire some prior knowledge that the steganographer adopts the channel-selected steganography. Then based on the χ^2 test (see Fig. 4 for instance), the probably-used channels can be predicted. Thus, it is proposed to utilize the Eq. (12) to deal with that problem of steganalysis. Besides, even if the steganographer adopts the traditional non-channel-selected steganography, our proposed adaptive detector can still perform well. Because in the establishment of the statistical model-based detector, some useless channels cannot contribute to our detection, instead might interfere with our detection. By effectively excluding the useless data, our designed adaptive detector improves its performance of detecting the channel-selected and non-channel-selected steganography.

In the other scenario, the steganalyzer might not know the used DCT channels for embedding. When assuming that the steganographer uses the channel-selected steganography, we can estimate the payload of embedding (or approximately predict the number of the used channels). It is feasible to estimate the DCT channels used for embedding with the help of the prior-art algorithms, such as [41], [42]. Then the design of the weight ω_l can refer to Eq. (12).

6.4 Adaptive Detector with Unknown Channel

Additionally, to further extend our adaptive steganalysis, when not assuming the proposed channel-selected steganography is utilized, it is proposed to assign the weights to each DCT channel as:

$$\omega_l = \frac{L+1-P_l}{\sum_{l=1}^{l=63} L+1-P_l} \tag{13}$$

where L denotes the number of the AC channels; P_l represents the position index of the channel in the ascend order acquired by the degree of model fitting (see details in Section 6.1). If the DCT channel preserves the high degree of model fitting, P_l with the large value corresponds to the backward position of the order. Because the steganographer less possibly embeds the secret information into the relevantly insecure channel. Therefore, it is proposed to assign the small weight to that channel. In fact, whether or not assuming that the steganographer utilizes the channel-selected steganography (or non-channel-selected one), the adaptive statistical model-based detector indeed improves its ability of classifying between cover and stego images, which can be empirically verified in our following experimental simulations.

7 Experimental Results

To empirically evaluate the reliability of current statistical mode-based detectors, we demonstrate the detection results of detecting non-channel-selected and the proposed channel-selected steganography. Then, to verify the effectiveness of the proposed adaptive statistical model-based detector, let us show its improved performance of detecting both steganographic schemes. Finally, compared with the prior-art detectors, the experimental results validate the better performance of our proposed adaptive detector.

7.1 Experimental Setup

Still, the benchmark BOSSbase dataset including 10000 uncompressed grey-level images with 512×512 pixels is used for our experiments. It is proposed to use 4 different quality factors and 6 different payloads to enrich the experimental dataset. For clarity, Table 3 illustrates the experimental image sets. Additionally, we adopt P_E, ROC, and Area Under Curve (AUC) as metrics.

TABLE 3: Experimental image sets.

Image source	BOSSbase 1.01 version [29]
Image color	Grey-level
Image size	512×512
Image format	JPEG
Quality factor	70, 85, 90, 95
Number of original images	10000
Payload	0.05, 0.1, 0.2, 0.3, 0.4, 0.5
Steganography method	Non-channel-selected and channel-selected, J-UNIWARD [9], UED [43], and UERD [44]
Non-adaptive steganalysis	Detectors of [3], [22], [23], [24], [25], [26], [32]
Adaptive steganalysis	Adaptive statistical model-based detector

7.2 Detection Performance of Non-adaptive Statistical Model-based Detector

First of all, we utilize the proposed channel-selected steganography (see Section 3.1 for details) to evaluate the reliability of current (or non-adaptive) statistical model-based detectors, respectively from [22], [25], and [26]. All 10000 images are compressed with quality factor 85, 90, and 95, serving as cover images. By using the channel-selected or non-channel-selected steganographic scheme, we can acquire the stego images with various payloads.

TABLE 4: Minimal P_E comparison, using the proposed channel-selected (CS) or non-channel-selected (Non-CS) steganographic algorithm (corresponding to the results of the second column with *italic* typeface); the statistical model-based detector using the algorithm of [25].

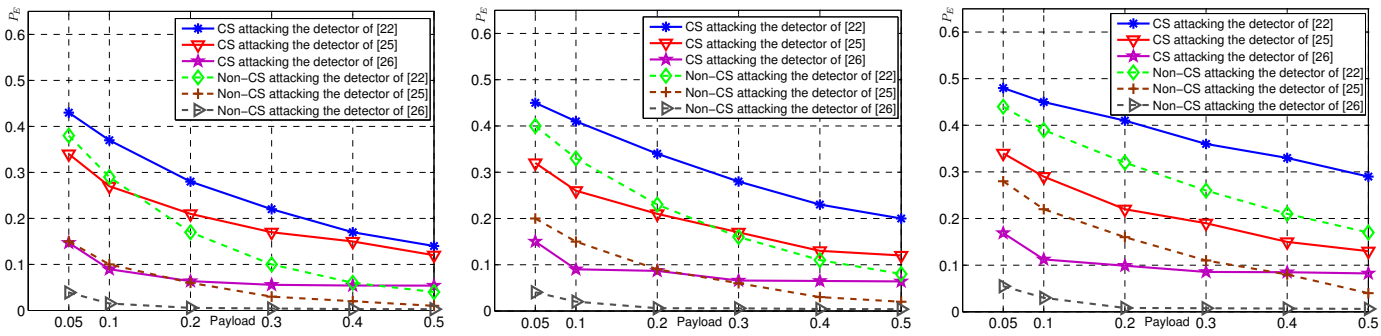
QF / Payload	85	90	95
0.05	0.34 *0.15*	0.32 *0.20*	0.34 *0.28*
0.10	0.27 *0.10*	0.26 *0.15*	0.29 *0.22*
0.20	0.21 *0.06*	0.21 *0.09*	0.22 *0.16*
0.30	0.17 *0.03*	0.17 *0.06*	0.19 *0.11*
0.40	0.15 *0.02*	0.13 *0.03*	0.15 *0.08*
0.50	0.12 *0.01*	0.12 *0.02*	0.13 *0.04*
Average	0.21 *0.06*	0.20 *0.09*	0.22 *0.15*

In Tables 4, and 5, we give the experimental results of the non-adaptive statistical model-based detectors with detecting both CS and Non-CS steganography. Note that, in Table 5, to specifically expose the slight differences with different payloads or QFs, we propose to retain the fourth decimal place.

The performances of the non-adaptive statistical detectors [25] and [26] are respectively illustrated in Tables 4 and 5. We have illustrated the results of the detector

1545-5971 (c) 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TDSC.2019.2962672, IEEE Transactions on Dependable and Secure Computing



(a) Quality factor 85 (b) Quality factor 90 (c) Quality factor 95

Fig. 5: Illustration of the minimal Prediction of Error (P_E); comparison from three different non-adaptive statistical detectors with steganalyzing JPEG images involving different quality factors: 85, 90, and 95. The plotted curves correspond to the results of Tables 2, 4, and 5.

TABLE 5: Minimal P_E comparison, using the proposed channel-selected (CS) or non-channel-selected (Non-CS) steganographic algorithm (corresponding to the results of the second column with *italic* typeface); the statistical model-based detector using the algorithm of [26].

QF / Payload	85	90	95
0.05	0.1469 *0.0389*	0.1501 *0.0405*	0.1689 *0.0548*
0.10	0.0897 *0.0151*	0.0901 *0.0201*	0.1123 *0.0299*
0.20	0.0638 *0.0056*	0.0867 *0.0060*	0.0988 *0.0081*
0.30	0.0557 *0.0043*	0.0658 *0.0058*	0.0854 *0.0076*
0.40	0.0541 *0.0030*	0.0649 *0.0042*	0.0849 *0.0069*
0.50	0.0539 *0.0028*	0.0639 *0.0038*	0.0822 *0.0062*
Average	0.0774 *0.0116*	0.0869 *0.0134*	0.1054 *0.0189*

[22] in Table 2 of Section 3. Similarly, when detecting the channel-selected steganography, the non-adaptive statistical detectors perform worse than the case with the non-channel-selected strategy. In the practical detection, the steganalysis detector proposed in [26] performs slightly better than that of [25], and considerably better than that of [22]. Because the detector of [26] adopts the mix model better describing the distribution than the counterpart from [22] or [25]. Note that the performance of the mix model-based detector in Fig. 1 is verified in Fig. 5 as well. Nevertheless, when dealing with channel-selected steganography, those non-adaptive detectors [22], [25], [26] cannot perform as well as the results of detecting non-channel-selected strategy, especially in that very common scenario that the payload is small and the quality factor is large.

To further clearly evaluate the compared performance of current statistical model-based detectors of detecting both steganographic schemes, let us illustrate the experimental results of Tables 2, 4, and 5 in Fig. 5. Whichever model-based paradigm such of [22], [25], [26] is adopted, compared to the non-channel-selected version, the channel-selected steganography indeed gives rise to larger P_E, implying that the unreliability of the non-adaptive statistical model-based detector.

7.3 Detection Performance of Adaptive Statistical Model-based Detector

One major contribution of this paper is that we devise the adaptive statistical model-based detector (see Eq. (11)). Note that we propose two types of adaptive detectors with known/unknown channel. By adopting different statistical models, we straightforward extend the adaptive detectors, and evaluate the performance of them. All 10000 images are compressed with quality factor 70 and 85. It is proposed to use the ROC curves to compare the performance among adaptive statistical detectors.

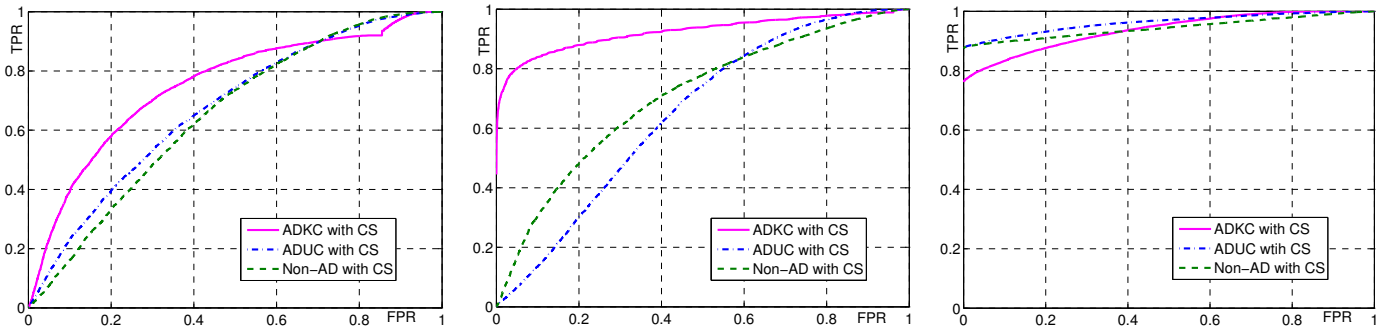
In the practical context, it should be noted that the adaptive statistical detectors are established based on the models respectively proposed in [22], [25], and [26]. We have empirically analyzed that the performance of the non-adaptive detectors of detecting the channel-selected steganography is degraded (see Tables 2, 4, and 5). Then, let us evaluate the performance of the adaptive detectors of detecting the CS/Non-CS steganography.

Fig. 6 and Fig. 7 illustrate the ROC curves of the statistical detectors for JPEG steganography, involving CS and Non-CS strategy. As Fig. 6(a) and Fig. 7(a) describe, the adaptive detector with known channel remarkably improves the detection performance when detecting the CS/Non-CS steganography. It is reasonable that the adaptive detector with knowing selected channels outperforms the one with unknown channel. Because the adaptive detector relying on all the channels with different weights, instead of predicting the most possibly-used channels, leading to that its detection performance approaches to the non-adaptive detector's results. Besides, when dealing with the case of Non-CS steganography, the proposed adaptive detector can still be effective, which indeed improves the detection accuracy of [22].

However, at the given low FPR, 0.1 for instance, using the model proposed by [22] (compared with two other models), the detection performances cannot achieve the satisfying results caused by the inaccuracy of model description. Besides, one can observe that with increasing the value of FPR (from around 0.805 to 0.835), the corresponding TPR continuously remains the same, leading to the abnormal distribution of the ROC curve. By investigating the LR

1545-5971 (c) 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TDSC.2019.2962672, IEEE Transactions on Dependable and Secure Computing

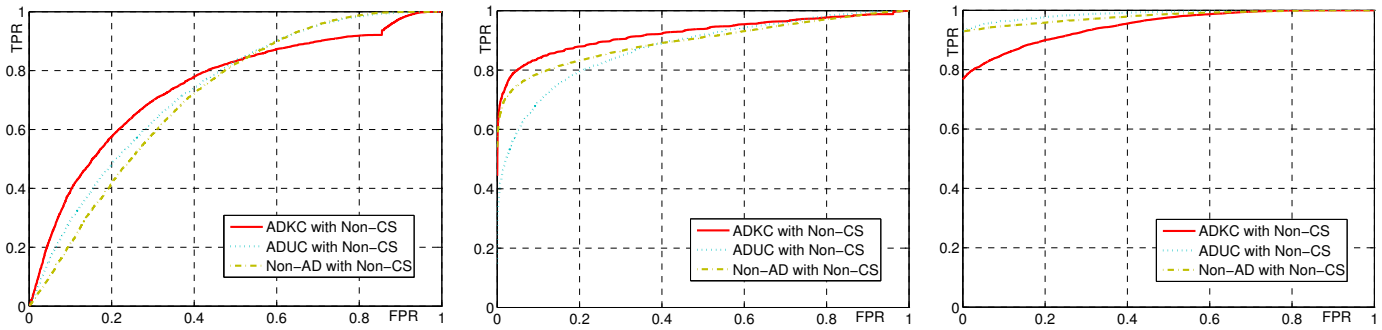


(a) Performance comparison with the case of quality factor 70 and payload 0.05; the statistical detectors are designed by using the model proposed in [22].

(b) Performance comparison with the case of quality factor 85 and payload 0.05; the statistical detectors are designed by using the model proposed in [25].

(c) Performance comparison with the case of quality factor 70 and payload 0.05; the statistical detectors are designed by using the model proposed in [26].

Fig. 6: Illustration of ROC curves; comparison from different statistical adaptive detectors of detecting CS steganography involving different quality factors: 70 and 85. ADKC represents adaptive detector with known channel for abbreviation; ADUC for adaptive detector with unknown channel; Non-AD for non-adaptive detector.



(a) Performance comparison with the case of quality factor 70 and payload 0.05; the statistical detectors are designed by using the model proposed in [22].

(b) Performance comparison with the case of quality factor 85 and payload 0.05; the statistical detectors are designed by using the model proposed in [25].

(c) Performance comparison with the case of quality factor 70 and payload 0.05; the statistical detectors are designed by using the model proposed in [26].

Fig. 7: Illustration of ROC curves; comparison from different statistical adaptive detectors of detecting Non-CS steganography involving different quality factors: 70 and 85.

$\widehat{\Lambda}_{apt}$ of all cover/stego images, we observe that the LR values of about 150 images from BOSSbase dataset equal to zero. Still, due to the inaccuracy of the model, the adaptive detector with limited selected channels becomes invalid when dealing with those images, regardless of CS/Non-CS steganography. Thus, the more accurate models such as [25] or [26] are required to help us improve the performance of the adaptive detector.

Using the model of [25], we compare the results from the statistical detectors in Fig. 6(b) and Fig. 7(b). Obviously, the CS steganography is capable of degrading the performance of the non-adaptive detector of [25]. However, by modifying the detector and directly selecting DCT channels (assigning binary weights, see Eq. (12)), the adaptive detector can effectively guarantee the high performance of detection in the cases of both CS and Non-CS steganography. Unfortunately, the result of the adaptive detector with unknown channel performs worst, implying that by assigning multiple weights (see Eq. (13)) to the channel based on the model of [25] cannot effectively detect the stego images. Because the proposed statistical model of [25] based on residual noise is relevantly sensitive to the channel change, leading to the sharply deceased detection rate.

The performance comparison from the statistical detectors designed by the model of [26] is demonstrated in Fig. 6(c) and Fig. 7(c). By considering the CS steganography, the reliability of the non-adaptive detector cannot be guaranteed. Through designing the adaptive detector with unknown channel, the detection performance (AUC, Area Under Curve) can be improved (see Table 6). However, at the given low FPR, the TPR of the non-adaptive detector still performs better than adaptive ones with known channel. Because if the description of statistical model is considerably accurate (better than the models respectively proposed by [22] and [25]), the designed adaptive detector with channel selection cannot provide a large scale of variables (of changed elements caused by data embedding), leading to the unsatisfying classification. In fact, in the case that the proposed statistical model can very accurately describe the distribution, we are prone to collect plenty of effective variables (DCT coefficients for instance) to establish our classifier, in which the differences caused by embedding

1545-5971 (c) 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TDSC.2019.2962672, IEEE Transactions on Dependable and Secure Computing

TABLE 6: Area Under Curve (AUC) comparison adopted by the statistical model [26] (ROC curves are illustrated in Fig. 6(c) and Fig. 7(c) as well), using the proposed channel-selected (CS) or non-channel-selected (Non-CS) steganographic algorithm.

Steganalysis \ Steganography	CS	Non-CS
Adaptive detector with unknown channel	**0.9621**	**0.9881**
Adaptive detector with known channel	0.9369	0.9490
Non-adaptive detector [26]	0.9446	0.9794

can be accumulated with limited errors. On the contrary, if the model is not very accurate (the model proposed in [22] for instance), the accumulated differences hardly help us design the well-performed statistical model-based detector with high detection rate; meanwhile the accumulated errors of parameter estimation (see our prior discussion in Section 5.2) probably offset the accumulated differences, which can also explain the abnormal results of the adaptive detector with known channel in Fig. 6(a) and Fig. 7(a). Nevertheless, when the large FPR (0.395 for CS steganography) is prescribed, our proposed adaptive detector with known channel outperforms the non-adaptive detector.

In the case that we adopt the accurate mix model proposed in [26], the adaptive detector with unknown channel even outperforms all other detectors (including adaptive ones with known channel), dealing with both CS and Non-CS steganography. The results are different from that of the adaptive detectors designed by using models of [22], [25], where the detectors with known channel perform better than the detectors with unknown channel.

In the practical steganalysis, we have to guarantee "enough" and "effective" DCT channels used for detection. However, when the statistical model is not very accurate such as the models of [22], [25], the problem of selecting between "enough" and "effective" DCT channels becomes a trade-off one. Then we give priority to "effective" DCT channels while to some extent sacrifice "enough" DCT channels. Due to the inaccuracy of the assumed model, the sum of likelihood ratio (see Eq. (11)) unavoidably contains some outliers, directly resulting in that "enough" DCT channels probably bring the accumulated prediction errors. Thus, we choose the ADKC (adaptive detector with known channel) to limit the accumulated prediction errors, which uses less coefficients than that of ADUC (adaptive detector with unknown channel).

On the contrary, when the statistical model is accurate enough such as the model of [26], we prefer to consider "enough" and "effective" DCT channels used for detection. The ADUC uses more coefficients than that of the ADKC. In this scenario, the very few of outliers hardly impacts the sum of likelihood ratio. The "enough" and "effective" DCT channels can be perfectly fitted by the assumed model; the sum of likelihood ratio calculated by a large number of effective DCT coefficients further magnifies the differences caused by embedding. Thus, the discrimination factor, referring to as the sum of likelihood ratio, becomes more discriminative, directly leading to the better-performed ADUC.

To further comprehensively demonstrate the detection performance of our proposed algorithm, it is proposed to compare a series of statistical model-based detectors. Specifically, we adopt three embedding schemes for comparing the minimal P_E from six different statistical model-based detectors, and the Avg. value by averaging P_E for each detector. It should be noted that 10000 cover images from the BOSSbase dataset are used for generating 10000 stego images.

TABLE 7: Minimal P_E comparison from a series of statistical model-based detectors, using embedding algorithms such as CS, Non-CS, and LSBM.

Steganalysis \ Steganography	CS	Non-CS	LSBM	Avg.
Model-based detector [32] designed for LSBM	0.4999	0.4996	0.4629	0.4875
Model-based detector [22]	0.3800	0.3400	0.4498	0.3899
Model-based detector [24]	0.4304	0.2679	0.4996	0.3993
Model-based detector [23]	0.4961	0.1364	0.4979	0.3768
Model-based detector [25]	0.4962	0.0562	0.4981	0.3502
Non-adaptive detector [26]	0.0620	0.0363	0.1812	0.0932
Adaptive detector with known channel	0.1170	0.1122	0.2363	0.1552
Adaptive detector with unknown channel	**0.0609**	**0.0357**	**0.1795**	**0.0920**

As Table 7 illustrates, our proposed adaptive detector with unknown channel (ADUC) performs best while the detector [32] gains the highest averaged P_E when testing three embedding schemes. Relying on the hypothesis testing theory, [22] first proposed to use the Laplacian model to detect Non-CS steganography. Next, by modifying the estimation algorithm of model parameter, both [24] and [23] further improved the detection accuracy for steganalysis. In [25], through further optimizing the statistical models and denoising filter used by [23], [24], the detection performance of the statistical model-based detector was advanced again. It is worth noting that when detecting LSB matching (randomly incrementing or decrementing LSB to match the secret bit to be embedded when necessary), our proposed adaptive detectors can still perform very well. Since the changes caused by LSBM in the spatial domain unavoidably give rise to the modification of DCT coefficients in the frequency domain, our proposed detector effectively captures the discriminative features based on the accurate statistical model.

In fact, the special detector[3] [32] was originally designed for dealing with the problem of LSBM. Thus, when detecting CS or Non-CS steganography, the detector [32] becomes invalid. Besides, due to the heterogeneous content of various images, the proposed model of [32] cannot very accurately describe all the images in the dataset. In this case, by increasing the payload, the more discriminative features caused by embedding can be acquired via the established statistical model-based detector. Thus, the detector [32] targeting stego images with the large payload hardly performs very well when testing a stego one with the small payload such as 0.05 in Table 7. For clear illustration, we also present the comparison results of Table 7 in Fig. 8.

Note that, relying on the mix model, our proposed adaptive detector with unknown channel is the optimal

3. The MATLAB source codes used for simulation are downloaded from the author's personal homepage: https://remi.cogranne.pagesperso-orange.fr.

1545-5971 (c) 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TDSC.2019.2962672, IEEE Transactions on Dependable and Secure Computing

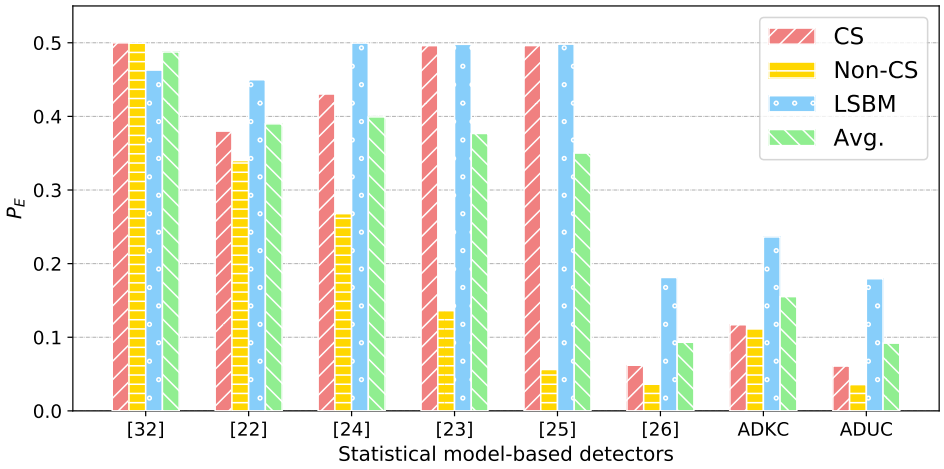


Fig. 8: Illustration of the minimal P_E comparison from a series of statistical model-based detectors, using embedding algorithms such as CS, Non-CS, and LSBM. The bars in Fig. 8 correspond to the results of Table 7.

choice of establishing the statistical model-based detector, which outperforms the detector proposed by [26]. Furthermore, under our proposed general framework, if the more accurate model is acquired, it is reasonable that the better-performed unsupervised detector can be designed. In fact, the performance of our unsupervised adaptive detector has outperformed that of the prior-art supervised detector. The specific discussion will be extended in the following subsection.

7.4 Comparison with Prior-art Supervised Detector

In our comparison experiments, all 10000 images are compressed with quality factor 70. Let us train the state-of-the-art ensemble classifier for steganalysis proposed in [3]. The feature extractor of [30] helps us obtain a high-dimensional vector of 22510 features. It should be noted that we randomly select 5000 image for training while the remaining images serve for testing. For fair comparison, when evaluating the performance of a supervised/unsupervised detector, we have to ensure that the same 5000 cover JPEG images, and their corresponding 5000 stego images by using steganographic strategies with payload 0.05, are used for detection. In addition, to generate the stego images including training and testing samples, we use the same steganography.

Tables 8 and 9 report the comparison results when we adopt CS and Non-CS steganography. The AUC and P_E serve as the comparison metrics. Obviously, it can be observed that when our proposed CS steganography is used, the performances of both ensemble and statistical detectors are degraded. Note that when dealing with the CS strategy, the performance of the supervised classifier [3] is the worst. To our knowledge, the ensemble classification algorithm nearly dominates the study of current steganalytic detectors, due to the high detection rate. However, without any training mechanism, our proposed adaptive statistical detector indeed outperforms that prior art [3] by considering the AUC or P_E results. Again, we need to emphasize that whichever steganographic algorithm (CS or Non-CS) is used, our designed adaptive statistical detector (based on the mix model) with unknown channel is the current optimal choice.

TABLE 8: Minimal P_E comparison, using the proposed channel-selected (CS) or non-channel-selected (Non-CS) steganographic algorithm.

Steganalysis \ Steganography	CS	Non-CS
Ensemble classifier [3] with ccJRM features [30]	0.2415	0.0481
Non-adaptive detector [26]	0.0803	0.0409
Adaptive detector with known channel	0.1392	0.1300
Adaptive detector with unknown channel	**0.0767**	**0.0389**

TABLE 9: Area Under Curve (AUC) comparison, using the proposed channel-selected (CS) or non-channel-selected (Non-CS) steganographic algorithm.

Steganalysis \ Steganography	CS	Non-CS
Ensemble classifier [3] with ccJRM features [30]	0.8481	0.9889
Non-adaptive detector [26]	0.9311	0.9787
Adaptive detector with known channel	0.9247	0.9418
Adaptive detector with unknown channel	**0.9567**	**0.9895**

Next, let us compare the detection performance of supervised and unsupervised detectors by using modern adaptive steganographic algorithms such as J-UNIWARD [9], UED [43], and UERD [44]. Under the framework of STC, three steganographic strategies aim at minimizing the distortion caused by embedding, leading to that the secret bits prefers hiding in the coefficients with low embedding cost. In this case, the statistical distribution of the image hardly can be accurately modelled, that indeed improves the undetectability of steganography. As Tables 10 and 11 illustrate, our proposed adaptive statistical model-based detectors cannot perform as well as the results of Tables 8 and 9. Because our proposed statistical model fails to describe the image containing the adaptive embedded bits by J-UNIWARD, UED, or UERD. As Tables 10 and 11 report, even though the powerful ensemble classifier [3] is adopted, the detection results are also close to random guess, which are very similar to that of our adaptive detectors.

1545-5971 (c) 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TDSC.2019.2962672, IEEE Transactions on Dependable and Secure Computing

TABLE 10: Minimal P_E comparison, using modern JPEG steganographic algorithms such as J-UNIWARD [9], UED [43], and UERD [44].

Steganalysis \ Steganography	J-UNIWARD	UED	UERD
Ensemble classifier [3] with ccJRM features [30]	**0.4880**	**0.4473**	**0.4771**
Non-adaptive detector [26]	0.4951	0.4895	0.4900
Adaptive detector with known channel	0.4970	0.4953	0.4968
Adaptive detector with unknown channel	0.4951	0.4900	0.4909

TABLE 11: Area Under Curve (AUC) comparison, using modern JPEG steganographic algorithms such as J-UNIWARD [9], UED [43], and UERD [44].

Steganalysis \ Steganography	J-UNIWARD	UED	UERD
Ensemble classifier [3] with ccJRM features [30]	**0.5147**	**0.5693**	**0.5292**
Non-adaptive detector [26]	0.5056	0.5151	0.5133
Adaptive detector with known channel	0.5012	0.5039	0.5015
Adaptive detector with unknown channel	0.5049	0.5130	0.5093

In fact, when the small payload such as 0.05 is used for modern adaptive steganography, current steganalysis cannot perform as well as the case of traditional non-adaptive steganography, that is the limitation of both supervised and unsupervised detectors.

8 Conclusion

In this paper, we propose the general framework of statistical model-based detectors, and the performance (involving model selection, parameters estimation, and payload prediction) is analyzed. To improve the reliability of the detectors, we propose two types of adaptive statistical model-based detectors, based on three distribution models. Experimental results verify that one of our designed adaptive detectors outperforms current unsupervised statistical model-based detectors such as [22], [25], [26], and also the supervised ensemble classifier [3], in the both cases of detecting channel-selected and non-channel-selected steganography. Furthermore, the proposed general framework can be extended to any design of adaptive model-based steganalytic detector. However, when steganalyzing stego images generated by modern adaptive steganographic algorithms with the small payload, both supervised and unsupervised detectors fail.

Currently, the TPR of few unsupervised detectors outperformed that of the supervised steganalysis designed by the ensemble classifier. However, our proposed adaptive detector outperforms the prior art, which indeed opens the alternative way of studying the adaptive detectors (not only focusing on the design of supervised classifiers based on the rich model features). In the future, we would like to continue our study focusing on the design of assigning more accurate weights over DCT channels for improving the detection accuracy of adaptive steganalysis.

9 Acknowledgements

This work was supported by the Natural Science Foundation of China under grant No. 61702150, U1804263, U1636219, U1536108, and U1736213, the Public Research Project of Zhejiang Province under grant No. LGG19F020015, the National Key R&D Program of China under grant No. 2016YFB0801303 and 2016QY01W0105, the Plan for Scientific Innovation Talent of Henan Province under grant No. 2018JR0018, the Key research and development plan project of Zhejiang Province under grant No. 2017C01062 and No.2017C01065.

References

[1] S. Lyu and H. Farid, "Steganalysis using higher-order image statistics," *Information Forensics and Security, IEEE Transactions on*, vol. 1, no. 1, pp. 111–119, 2006.

[2] X. Luo, F. Liu, S. Lian, C. Yang, and S. Gritzalis, "On the typical statistic features for image blind steganalysis," *IEEE Journal on selected areas in Communications*, vol. 29, no. 7, pp. 1404–1422, 2011.

[3] J. Kodovskỳ, J. Fridrich, and V. Holub, "Ensemble classifiers for steganalysis of digital media," *Information Forensics and Security, IEEE Transactions on*, vol. 7, no. 2, pp. 432–444, 2012.

[4] J. Ye, J. Ni, and Y. Yi, "Deep learning hierarchical representations for image steganalysis," *Information Forensics and Security, IEEE Transactions on*, vol. 12, no. 11, pp. 2545–2557, 2017.

[5] C. F. Tsang and J. Fridrich, "Steganalyzing images of arbitrary size with cnns," *Electronic Imaging*, vol. 2018, no. 7, pp. 1–8, 2018.

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

[7] 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.

[8] 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.

[9] 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, pp. 1–13, 2014.

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

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

[12] X. F. Song, F. Liu, C. Yang, X. Luo, and Y. Zhang, "Steganalysis of adaptive jpeg steganography using 2d gabor filters," in *Proceedings of the 3rd ACM Workshop on Information Hiding and Multimedia Security*. ACM, 2015, pp. 15–23.

[13] Y. Ma, X. Luo, X. Li, Z. Bao, and Y. Zhang, "Selection of rich model steganalysis features based on decision rough set α-positive region reduction," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 29, no. 2, pp. 336–350, 2019.

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

[15] J. Yang, K. Liu, X. Kang, E. Wong, and Y. Shi, "Steganalysis based on awareness of selection-channel and deep learning," in *International Workshop on Digital Watermarking*. Springer, 2017, pp. 263–272.

[16] A. D. Ker, P. Bas, R. Böhme, R. Cogranne, S. Craver, T. Filler, J. Fridrich, and T. Pevnỳ, "Moving steganography and steganalysis from the laboratory into the real world," in *Proceedings of the first ACM workshop on Information Hiding and Multimedia Security*. ACM, 2013, pp. 45–58.

[17] J. Fridrich and M. Goljan, "On estimation of secret message length in lsb steganography in spatial domain," in *Security, Steganography, and Watermarking of Multimedia Contents VI*, vol. 5306. International Society for Optics and Photonics, 2004, pp. 23–35.

[18] O. Dabeer, K. Sullivan, U. Madhow, S. Chandrasekaran, and B. Manjunath, "Detection of hiding in the least significant bit," *Signal Processing, IEEE Transactions on*, vol. 52, no. 10, pp. 3046–3058, 2004.

[19] A. D. Ker and R. Böhme, "Revisiting weighted stego-image steganalysis," in *Electronic Imaging 2008*. International Society for Optics and Photonics, 2008, pp. 681 905–681 905.

[20] R. Böhme, "Weighted stego-image steganalysis for jpeg covers," in *International Workshop on Information Hiding*. Springer, 2008, pp. 178–194.

1545-5971 (c) 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TDSC.2019.2962672, IEEE Transactions on Dependable and Secure Computing

[21] R. Cogranne, C. Zitzmann, F. Retraint, I. V. Nikiforov, P. Cornu, and L. Fillatre, "A local adaptive model of natural images for almost optimal detection of hidden data," *Signal Processing*, vol. 100, pp. 169–185, 2014.

[22] C. Zitzmann, R. Cogranne, L. Fillatre, I. Nikiforov, F. Retraint, and P. Cornu, "Hidden information detection based on quantized laplacian distribution," in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*. IEEE, 2012, pp. 1793–1796.

[23] T. Qiao, C. Zitzmann, R. Cogranne, and F. Retraint, "Detection of jsteg algorithm using hypothesis testing theory and a statistical model with nuisance parameters," in *Proceedings of the 2nd ACM Workshop on Information Hiding and Multimedia Security*. ACM, 2014, pp. 3–13.

[24] T. Qiao, C. Zitzmann, F. Retraint, and R. Cogranne, "Statistical detection of jsteg steganography using hypothesis testing theory," in *Image Processing (ICIP), 2014 IEEE International Conference on*. IEEE, 2014, pp. 5517–5521.

[25] T. Qiao, F. Retraint, R. Cogranne, and C. Zitzmann, "Steganalysis of jsteg algorithm using hypothesis testing theory," *EURASIP Journal on Information Security*, vol. 2015, no. 1, pp. 1–16, 2015.

[26] T. H. Thai, R. Cogranne, and F. Retraint, "Statistical model of quantized dct coefficients: application in the steganalysis of jsteg algorithm," *Image Processing, IEEE Transactions on*, vol. 23, no. 5, pp. 1980 – 1993, 2014.

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

[28] D. Upham, "Jsteg steganographic algorithm," *Available on the Internet http://www.filewatcher.com/m/jpeg-jsteg-v4.diff.gz.8878-0.html*, 1999.

[29] 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.

[30] J. Kodovskỳ and J. Fridrich, "Steganalysis of jpeg images using rich models," in *Media Watermarking, Security, and Forensics 2012*, vol. 8303. International Society for Optics and Photonics, 2012, pp. 1–13.

[31] T. H. Thai, R. Cogranne, and F. Retraint, "Optimal detection of outguess using an accurate model of dct coefficients," in *Information Forensics and Security (WIFS), 2014 IEEE 6th International Workshop on*, 2014, pp. 1785–1790.

[32] R. Cogranne and F. Retraint, "An asymptotically uniformly most powerful test for lsb matching detection," *Information Forensics and Security, IEEE Transactions on*, vol. 8, no. 3, pp. 464–476, 2013.

[33] E. L. Lehmann and J. P. Romano, "Testing statistical hypotheses," in *Second Edition*. Springer, 2005.

[34] C. Zitzmann, R. Cogranne, F. Retraint, I. Nikiforov, L. Fillatre, and P. Cornu, "Statistical decision methods in hidden information detection," in *International Workshop on Information Hiding*. Springer, 2011, pp. 163 – 177.

[35] F. Muller, "Distribution shape of two-dimensional dct coefficients of natural images," *Electronics Letters*, vol. 29, no. 22, pp. 1935–1936, 1993.

[36] J. H. Chang, J. W. Shin, N. S. Kim, and S. K. Mitra, "Image probability distribution based on generalized gamma function," *Signal Processing Letters, IEEE*, vol. 12, no. 4, pp. 325–328, 2005.

[37] P. Sallee, "Model-based methods for steganography and steganalysis," *International Journal of Image and Graphics*, vol. 5, no. 01, pp. 167–189, 2005.

[38] R. Böhme and A. Westfeld, "Breaking cauchy model-based jpeg steganography with first order statistics," in *Computer Security–ESORICS 2004*. Springer, 2004, pp. 125–140.

[39] E. Y. Lam and J. W. Goodman, "A mathematical analysis of the dct coefficient distributions for images," *Image Processing, IEEE Transactions on*, vol. 9, no. 10, pp. 1661–1666, 2000.

[40] J. Kodovsky and J. Fridrich, "Quantitative structural steganalysis of jsteg," *Information Forensics and Security, IEEE Transactions on*, vol. 5, no. 4, pp. 681–693, 2010.

[41] A. D. Ker, "Locating steganographic payload via ws residuals," in *Proceedings of the 10th ACM workshop on Multimedia and security*. ACM, 2008, pp. 27–32.

[42] T.-T. Quach, "Optimal cover estimation methods and steganographic payload location," *Information Forensics and Security, IEEE Transactions on*, vol. 6, no. 4, pp. 1214–1222, 2011.

[43] L. Guo, J. Ni, and Y. Q. Shi, "Uniform embedding for efficient jpeg steganography," *Information Forensics and Security, IEEE Transactions on*, vol. 9, no. 5, pp. 814–825, 2014.

[44] L. Guo, J. Ni, W. Su, C. Tang, and Y. Q. Shi, "Using statistical image model for jpeg steganography: uniform embedding revisited," *Information Forensics and Security, IEEE Transactions on*, vol. 10, no. 12, pp. 2669–2680, 2015.

Tong Qiao received the B.S. degree in Electronic and Information Engineering in 2009 from Information Engineering University, Zhengzhou, China, and the M.S. degree in Communication and Information System in 2012 from Shanghai University, Shanghai, China, and the Ph.D. degree in University of Technology of Troyes, Laboratory of Systems Modelling and Dependability, Troyes, France. Since 2016, he works as an assistant professor in School of Cyberspace from Hangzhou Dianzi University. His current research interests focus on steganalysis and digital image forensics.

Xiangyang Luo received his B.S., M.S., and Ph.D degrees from the State Key Laboratory of Mathematical Engineering and Advanced Computing, Zhengzhou, China, in 2001, 2004, and 2010, respectively. He is the author or co-author of more than 100 refereed international journal and conference papers. He is currently a Professor of the State Key Laboratory of Mathematical Engineering and Advanced Computing. His research interests are image steganalysis and Forensics.

Ting Wu received the Ph.D. degree from Shandong University, in 2002. He is currently a Professor of Hangzhou Dianzi University. His research interest is cryptography.

Ming Xu received his M.S., and Ph.D. degrees from Zhejiang University, in 2000, and 2004, respectively. He is currently a Professor of Hangzhou Dianzi University. His research interest is digital forensics.

Zhenxing Qian received the B.S. and Ph.D. degrees from the University of Science and Technology of China (USTC) in 2003 and 2007, respectively. He is currently a Professor with the School of Computer Science, Fudan University. He has published over 100 peer-reviewed papers on international journals and conferences. His research interests include information hiding, image processing, and multimedia security.

1545-5971 (c) 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.