RESEARCH ARTICLE

HYSA: HYbrid Steganographic Approach using multiple steganography methods

Hedieh Sajedi* and Mansour Jamzad

Department of Computer Engineering, Sharif University of Technology, Tehran, Iran

ABSTRACT

The aim of image steganography is to hide data in an image with visual and statistical invisibility. There have been a number of steganography techniques proposed over the past few years. In sequence, the development of these methods has led to an increased interest in steganalysis techniques. Most of the steganalysis methods attempt to estimate cover image statistics. One way to provide a secure steganography method is to disturb the estimation of steganalyzers. The main concern of this paper is to resist against steganalysis methods and utilize a mechanism to securely embed more secret data into an image. We present HYbrid steganographic approach (HYSA), which embeds secret data with randomized embedding algorithms in randomized regions of the cover image. HYSA embeds secret data in only some regions of the image not all over the image. In addition, HYSA can employ more than one steganography method for embedding secret data in the image. Random selection of embedding regions and steganography methods leaves a combination of several types of distortion on the image, which are difficult to be recognized by the steganalysis methods. Experimental results by applying some well-known and efficient steganalysis methods illustrate that the proposed steganographic approach is more undetectable than some popular steganography methods. Copyright © 2010 John Wiley & Sons, Ltd.

KEYWORDS

covert communication; information hiding; randomized hiding; steganalysis; steganography

*Correspondence

Hedieh Sajedi, Department of Computer Engineering, Sharif University of Technology, Tehran, Iran.
E-mail: a_sajedi@ce.sharif.edu

1. INTRODUCTION

The primary goal of steganography is to hide a covert communication within an innocuous communication. The main requirement of steganography is undetectability, which means that it could not be determined whether an object contains hidden data. Research in steganalysis, the science of detecting the presence of hidden data in an innocuous-looking object, has taken great advances in the past few years. Statistical steganalysis methods work by evaluating a suspected stego object against an assumed or computed cover distribution or model. Blind statistical steganalysis methods use a supervised machine learning technique trained on features derived from cover as well as stego objects. This class of steganalysis methods has been very successful in detecting steganography methods available today. Detection results presented in References [1,2] indicate that popular steganography methods, such as Outguess [3], Steghide [4], Model-based steganography [5], and 1D statistical restoration methods [6,7], can now be detected using some blind steganalysis techniques such as those given in References [1,8–13].

To provide secure communication in the presence of steganalysis, the steganographer must embed data into the cover image in such a way that no image features are significantly perturbed during the embedding process. Most steganography methods can avoid being detected by current steganalysis techniques simply by decreasing the amount of data embedded in a cover image. Most of the steganalyzers depend on the stego image to derive the approximate cover image statistics \textit{via} some sort of self-calibration process [2]. Instead of (or along with) trying to preserve the feature vectors, the steganographer can embed data in such a way that it distorts the steganalyzers’ estimate of the cover image statistics. Sarkar \textit{et al.} in Reference [2] represent that estimation distortion of steganalyzers can be done by randomizing the hiding process.

In this paper, we employ the approach of randomized hiding to disturb the statistical features computed by the blind steganalysis methods and propose HYbrid steganographic approach (HYSA), which is a twin-randomized solution.
In HYSA, first, the steganographer randomly chooses one or more non-overlapping regions (i.e., a rectangular area) of the cover image and then he selects one or more predefined steganography methods in random. For simplicity, we consider the number of random-selected regions and steganography methods the same. Subsequently, the secret data are embedded in the chosen regions with the selected steganography methods. In this approach, the location and size of the embedding regions and the name of the steganography method corresponding to each region are the embedding parameters that should be determined. The intended receiver ought to be aware of the embedding parameters to extract the secret data correctly.

Usually the steganalysis methods assume that the secret data are embedded all over the image and they normalize the feature values based on the number of Discrete Cosine Transform (DCT) blocks or wavelet subband blocks. Therefore, if the steganographer selects some random regions in the cover image and embed only in those regions, the estimations of steganalyzers are defeated and the chance to detect the stego image is reduced. Furthermore, different steganography methods have dissimilar embedding procedures and each method affects on some particular statistical features of images. Therefore, employing HYSA may cause irregular and hard to detect statistical artifacts compared to utilizing only one embedding method in the image monotonously.

The novelty of this paper is proposing a hybrid steganography approach that attempts to disturb estimation of steganalyzers by leaving unknown and undetectable traces on stego images. This is obtained by two-phase randomization in the embedding process.

Our experimental results show that using HYSA, the rate of embedding measured in bits has been improved while maintaining the undetectability of produced stego images against recent blind steganalysis techniques.

This paper is organized as follows. In the next section, a background of the performed investigations is presented. In Section 3, we introduce and describe our hybrid steganography approach: HYSA. Experiments and validations of HYSA are presented in Section 4 and finally Section 5 concludes this paper.

2. BACKGROUND

Steganography approaches are often classified based on the steganographic cover type into image, audio, graph, game [14], text [15], etc. Most of the published steganography approaches hide secret data as noise in a cover object that is assumed to look innocent. For example, hiding a message in a text-cover by modifying the format and style of an existing text. However, such alteration of covers may raise suspicion [15].

Recently, noiseless steganography paradigm (Nostega) [16] has been proposed which embeds secret data in the cover rather than noise. Some methods have been developed based on the Nostega paradigm. One of them is summarization-based steganography [17] that exploits automatic summarization techniques to hide data in the auto-generated summary-cover (text-cover) that looks like an ordinary and legitimate summary. Another steganography method that is based on Nostega paradigm is list-based steganography (Listega) [15]. Listega takes advantage of textual list (e.g., products, subjects, books, etc.) to camouflage data by exploiting itemized data to conceal data. It encodes a message then assigns it to some legitimate items in order to generate a text-cover in the form of a list. The generated list of items can be embedded among other legitimate non-coded items for more protection. Mature linguistic steganography [18] exploits template techniques along with random series values (RS) to hide data without generating any suspicious pattern.

In graph steganography (Graphstega) [19], unlike other schemes, the secret data are naturally embedded in the cover by simply generating the cover based on the secret data. Graphstega hides secret data, as data points in a graph, and thus the secret data would not be detectable as noise. Similarly, Chestega [14] exploits popular games, such as chess, checkers, crosswords, domino, etc., to hide secret data.

The goal of steganography is to conceal the covert communication. Thus, a stego image containing a secret data should be inconspicuous to an adversary. Steganography and steganalysis methods are concerned with those properties of the stego image that can be quantified. Histograms and a variety of higher-order statistics are usually computed by steganalyzers to evaluate the compliancy of a given image. The presence of statistical anomalies is assumed to be indicative of non-legitimate use and, therefore, may be used by the adversary to decide that a covert communication is present or at least examine the stego image in more details.

All image steganography techniques utilize redundancies in cover images for embedding. However, they differ on their approach and the image format (e.g., BMP, JPEG) they work on. For example, techniques, such as F5 [20], outguess, model-based steganography, and perturbed quantization (PQ) [21], work on JPEG images by modifying the DCT coefficients. Although changing DCT coefficients will cause unnoticeable visual artifacts, they may impose detectable statistical changes. Some other methods embed data in wavelet [22,23] or contourlet [24] coefficients of image decomposition.

The effect of applying some steganography algorithms on different statistical models of natural images has been studied in Reference [25]. It shows that the statistics of natural images are altered by hiding non-natural data in them. For the class of natural images, the change generally falls within the intrinsic variability of the statistics, and thus does not allow for reliable detection, unless knowledge of the data hiding process is taken into account.

Since the statistical changes are used by steganalysis techniques to detect any embedded information, each of the mentioned steganography techniques tries to minimize the statistical changes with different approaches. In the sequel, we review some steganography methods briefly.
Model-based steganography (MB) technique models some of the statistical properties of an image and tries to preserve these models in the embedding process. MB method breaks down transformed image coefficients into two parts and replaces the perceptually insignificant component with the coded secret data. Initially, the marginal statistics of quantized non-zero AC DCT coefficients are modeled with a parametric density function. For this, a histogram of each frequency channel is obtained and the model is fit to each histogram by determining the corresponding model parameters. MB defines the offset value of coefficients within a histogram bin as a symbol and computes the corresponding symbol probabilities from the relative frequencies of symbols. In the heart of the embedding operation is a non-adaptive arithmetic decoder, which takes as input the secret data and decodes them with respect to measured symbol probabilities. Then, the entropy-decoded data are embedded by specifying new bin offsets for each coefficient. In other words, the coefficients in each histogram bin are modified with respect to embedding rule, while the global histogram and symbol probabilities are preserved.

The research in Reference [21] proposes PQ steganography method, in which the cover image goes through an information reducing operation and the secret data are embedded in this process. In PQ method, a JPEG image is recompressed with a lower quality factor, but at the quantization step in the compression process, DCT coefficients that could be quantized to an alternate bin with an error smaller than some set value are modified. At the detector, in order to determine which DCT coefficients carry the payload, Fridrich et al. [21] pose the problem in terms of writing into memory with defective cells.

YASS steganography method embeds data in image blocks whose locations are randomized. Some locations in an image are possible to hold an entire embedding block and some locations are definitely not. Additionally, YASS employs a Quantization Index Modulation embedding strategy in order to enhance the robustness of the embedded data [26].

In Reference [24], embedding is done in contourlet transform domain. In Contourlet-based steganography method, contourlet transform is applied to capture significant image coefficients across spatial and directional resolutions. Contourlet-based method embeds the secret data in some proper contourlet coefficients of the cover image. The embedding algorithm takes advantage of adaptive methods by embedding data in non-smooth regions and the procedure is carried on by changing the value of two contourlet coefficients to represent 0 or 1.

3. HYBRID STEGANOGRAPHIC APPROACH: HYSA

HYSA is proposed based on three concepts described in the following:

(1) The basis of steganalysis is on detecting differences between the images before and after embedding. The steganography procedure may change the continuity of a cover image. Consequently, it may reduce the correlation among adjacent pixels, bit-planes, and image blocks [10]. Each steganography method embeds in images in the same way and produces a particular type of distortion on stego images. Therefore, discovering the distortion type of a steganography method, which is the difference of some statistical characteristics between the cover and stego images, is the key issue in image steganalysis. If we disturb the differences by some random operations or produce inhomogeneity in embedding process, the steganalyzers could get confused and thus higher security for stego images is achieved. Therefore, in HYSA we force randomness in embedding process. Although HYSA reduces the continuity of cover images, but the types of distortions in different stego images differ from each other due to the randomized embedding. Hence, the continuity breaking is not detectable.

(2) The size of images affects on the performance of steganography or conversely the steganalysis methods. This aspect has been investigated from the practical point of view in the literature [2,27]. In addition, there have been a few theoretical studies exploring the relationship between dimension of the data and the steganalytic detection rate [28,29]. In Reference [28], Wang et al. use Kobayashi-Thomas bound in context of optimal statistical steganalysis to infer that the detector error for i.i.d. (independent and identically-distributed) data is greater than $\exp(-N)$, where $N$ is the dimensionality of the covertext or stegotext. Ker et al. in Reference [29] illustrate that the steganography capacity of a single cover image under some conditions is proportional to the square root of the cover image size. The work in Reference [27] has analyzed the impacts of the cover image size on the performance of steganography and steganalysis methods. Its results show that the detection performance of steganalysis methods is likely to suffer for smaller images, as the distinctiveness of the collected statistics is reduced [27]. In addition, Sarkar et al. in Reference [2] has mentioned that it is harder to detect smaller stego images. Considering the mentioned researches, it can be deduced that the statistical changes in small stego images are less detectable than that of large stego images. Therefore, instead of embedding in the entire of an image, HYSA embeds in some small sub-images of the image. Thus, instead of a large size stego image, HYSA produces smaller size stego sub-images, which are placed on the cover image. In this way, HYSA embedding may be less detectable than traditional embedding in whole of the image. Our experimental results in subsection 4.2 show that the distinctiveness of collected statistics of stego images produced by HYSA is reduced.
(3) Since different steganography methods apply dissimilar embedding operations for hiding one bit of secret data, each method may affect on particular statistical features of stego images. Therefore, if HYSA employs multiple steganography methods, it may cause irregular and unknown statistical artifacts compared to utilizing a single steganography method. Consequently, detection of stego images produced by HYSA is more difficult than detecting stego images produced by a single steganography method.

Employing HYSA, the steganographer selects some embedding regions \(ER = \{R_k | k = 1...K, \bigcap_{k=1}^{K} R_k = \emptyset\}\) of the cover image in random in such a way that there is no overlapping between them and \(R_k = \{(x, y)|a_k \leq x \leq a_k + H_k, b_k \leq y \leq b_k + W_k\}\) is a rectangle in the xy-plane of image. Then, he uses different randomly selected steganography methods \(S = \{S_k | k = 1...K\}\) to hide parts of secret data in the chosen regions. Each embedding in HYSA has some parameters that are location \((a_k, b_k)\), size (i.e., height \(H_k\) and width \(W_k\)), and the name of utilized steganography method \(S_k\) in each region.

HYSA approach needs sharing embedding parameters (i.e., \(a_k, b_k, H_k, W_k, S_k, k = 1,...,K\)) between sender and receiver. Traditional embedding mechanisms may require shared parameters (such as stego keys) and thus HYSA may be not much different in this respect but it increases the amount of shared information that must be coordinated.

The strategy of dividing the secret data to some parts depends on the number of selected regions denoted with \(K\). The secret data can be simply divided into \(K\) equal size parts if the size of embedding regions is equal. For regions with different sizes, the secret data can be divided to \(K\) parts, proportional to the size of regions. After determining the embedding parameters, the first part of secret data is embedded in the first selected region using the first selected steganography algorithm. Afterward the remaining non-hidden data are hidden in other selected regions with other selected steganography methods.

### 3.1. Embedding secret data

In HYSA, each steganography method hides a part of secret data in the cover image as it does in usual embedding. For example if the steganographer wants to hide data in an image using PQ and Contourlet-based steganography methods, he hides one part of data in DCT coefficients of a selected embedding region and the other part in contourlet coefficients of another selected region.

Figure 1 presents some samples of embedding schemes in HYSA. In scheme no. 1 (partial embedding), the steganographer randomly chooses a region \((K = 1)\) of the cover image for embedding. As we will see later in the Section 4, to hide a certain secret data, higher security can be obtained by partial embedding rather than full embedding (embedding in whole of the image). In scheme no. 2, the steganographer selects two regions \((K = 2)\) of the cover image and embeds the secret data using MB and PQ steganography methods. In scheme no. 3, he picks three regions \((K = 3)\) of the cover image randomly and then, the secret data are divided into three parts. One part of secret data is embedded in DCT domain using MB method, another part of the secret data is embedded in contourlet domain applying Contourlet-based method, and the remaining part is embedded in DCT domain utilizing PQ method.

Scheme no. 4 is similar to scheme no. 3, but instead of Contourlet-based steganography method, YASS is used to hide a part of secret data. In scheme no. 5, PQ, YASS, and Contourlet-based steganography methods are employed to conceal secret data.

We visualized only few samples of embedding schemes in Figure 1. However, the steganographer can employ a customized embedding scheme. The candidate regions for embedding can be selected previously, randomly, or based on the content of the cover image. For example, we can define ‘interesting regions for embedding’, identify them, and embed only in these regions. The main point is that the selected regions for embedding should be smaller than the original cover image and they should not have any overlaps. An optimization method can be applied to find the best embedding scheme and minimize the risk of detection by steganalyzers. It should be noted that all the steganography methods with different embedding domains could be participated in the embedding schemes of HYSA.

The steganographer is supposed to prepare the cover image and its selected regions with the format required for each steganography method. Each selected region is detached from the cover image, a part of secret data is embedded in it, and then its stego version is attached to the image in its place. For example in scheme no. 3, the output of PQ, MB, and Contourlet-based steganography methods should be three JPEG sub-images with the same quality factor. To obtain a stego image with quality factor of 70 (the original cover image is in raw format), a copy of cover image is compressed to JPEG format with quality factor of 70. Afterward, the steganographer detaches three regions of the original cover image in its raw format. He compresses one of the regions to JPEG format with quality factor of 70 and gives it to MB method as the cover sub-image. He alters another region to JPEG format with quality factor of 70 and gives it to Contourlet-based method. The remaining region is changed to JPEG format with quality factor of 85 and it is given to PQ method. As we know, in PQ steganography method, the cover images go through a recompression operation and the secret data are embedded in this process. Thus, the stego image produced by PQ method is a recompressed JPEG image (in our experiments to have a stego sub-image with quality factor of 70, the cover sub-image that is the input of PQ steganography method should have quality factor of 85). Finally, all the three stego sub-images are replaced in their places on the compressed version of the cover image. In this way, the stego image is a compressed version of the original cover image with three embedded stego sub-images.
Hybrid steganographic approach

Figure 1. Some samples of embedding schemes using model-based, perturb quantization, Contourlet-based, and YASS steganography methods.

Figure 2 is an illustration of embedding scheme no. 3 and the format of cover and stego images using MB, PQ, and Contourlet-based steganography methods.

The required time for embedding in HYSA is as that given in Equation (1).

\[ t = t_{ms} + t_{rd} + \sum_{k=1}^{K} t_{te} + t_{ta} \]  

where \( t \) is the total time of embedding in HYSA, \( t_{ms} \) is the time for steganography method selection, \( t_{rd} \) is the embedding regions detachment time, \( t_{te} \) is the embedding time of the \( k \)th steganography method, \( K \) is the number of selected embedding regions, and \( t_{ta} \) is the time of embedding regions attachment.

3.2. Extraction of hidden data

The receiver of stego images, knowing the embedding scheme and its parameters extracts each part of hidden data separately and then joins the extracted parts of secret data.
Hybrid steganographic approach

4. EXPERIMENTS

4.1. Experimental setups

Several simulations are performed to assess the performance of the proposed approach. Experiments are performed with 3164 images from camera image database of Binghamton University. Images are grayscale of size 512 × 512. All the images have been stored in a raw format (tif). These images were previously used for steganography in References [30,31].

To make stego datasets with various payloads, random binary data with different sizes are embedded into images using PQ, MB, Contourlet-based, and YASS steganography methods. We changed the values of the parameters in steganography methods several times to obtain stego image datasets with certain payloads. Each distinct stego image dataset with certain payload has 3164 stego images. To have uniformity with the steganography methods that work in other domains except DCT, we express embedding rate in bits. Every steganalyzer is achieved as follows: from the constructed model, 4000 images (2000 clean and 2000 stego) are employed to train a support vector machine (SVM). We then utilize the obtained classifier, to distinguish between cover and stego images in the test dataset. The test dataset has 2328 images (1164 clean and 1164 stego).

In our experiments, we divide the secret data equally between the applied steganography methods and we select equal size regions in the cover images for embedding. For steganalysis of PQ steganography method, we tried to distinguish between cover images with JPEG format (not re-compressed) and stego images in JPEG format with the same quality factor.

Security of HYSA is evaluated by the following efficient and famous steganalysis methods:

1. Wavelet-based steganalysis method (WBS) [9] builds a model for clean images by using higher-order statistics and then shows the deviation of stego images from the constructed model. Quadratic mirror filters (QMF) are used to decompose the image into wavelet domain, after which higher-order statistics, such as mean, variance, skewness, and kurtosis, are calculated for each subband, as the first part of the feature set. The second part includes the same statistics, which are calculated for the error obtained from an optimal linear predictor of coefficients’ magnitudes in each subband.

2. Markov-DCT-based steganalysis method (274-dim) has a 274-dimensional feature vector that merges Markov and DCT features of an image [1]. The merged DCT and Markov features have also been found to be among the most successful steganalysis methods for detecting image steganography.

3. The 324-dimensional feature vector steganalysis method (324-dim) proposed by Chen et al. [12] is an improvement of the 39-dimensional feature vector method [10]. This universal steganalysis method is based on statistical moments derived from both image 2-D array and JPEG 2-D array. In addition to the first-order histogram, the second-order histogram is considered. Consequently, the moments of 2-D characteristic functions are also included for steganalysis.

The steganalysis performance is quantified through the detection accuracy [26]. The SVM classifier has to distinguish between two classes of images: cover (class ‘0’) and stego (class ‘1’). Let \( X_0 \) and \( X_1 \) denote the events that the image being observed belongs to classes ‘0’ and ‘1’, respectively. On the detection side, let \( Y_0 \) and \( Y_1 \) denote the events that the observed image is classified as belonging to classes ‘0’ and ‘1’, respectively. We use detection accuracy \( D_a \) which is the per cent of detection probability \( P_d \) as our evaluation criteria according to the following equations:

\[
D_a = P_d \times 100
\]

\[
P_d = 1 - P_{\text{error}}
\]

\[
P_{\text{error}} = P(X_0)P(Y_1|X_0) + P(X_1)P(Y_0|X_1)
\]

\[
= \frac{1}{2}P_{\text{FA}} + \frac{1}{2}P_{\text{miss}} \quad \text{for} \ P(X_0) = P(X_1) = \frac{1}{2}\tag{2}
\]

where \( P_{\text{FA}} = P(Y_1|X_0) \) and \( P_{\text{miss}} = P(Y_0|X_1) \) denote the probability of false alarm and missed detection, respectively. The above equation assumes an equal number of cover and stego images in the dataset.

An uninformed detector can classify all the test images as stego (or cover) and get an accuracy of 50. Thus, \( D_a \) being close to 50 implies nearly undetectable hiding, and as the detectability improves, \( D_a \) should increase towards 100.

4.2. Performance evaluation of HYSA

Table I shows the detection accuracy of three steganalyzers in detection of stego images produced by HYSA using different embedding schemes. As we see, scheme no. 1 (partial embedding) provides higher security than the classical embedding using a single steganography method in entire of the cover image (full embedding). In addition, the results show that using combinations of different steganography methods (scheme no. 2–5) results in more secure stego images. In scheme no. 3–5 that three steganography techniques are used, we can divide the secret data into three parts of equal or unequal sizes. In this experiment, we selected three equal size embedding regions. For example, 6000 bits of secret data is divided into three parts of size 2000 bits. The accuracy of steganalysis methods in detection of stego images is calculated.

 DOI: 10.1002/sec

Table 1. Accuracy of steganalysis methods in detection of stego images produced by HYSA using different embedding schemes.

<table>
<thead>
<tr>
<th>Steganography method</th>
<th>Payload (bits)</th>
<th>WBS</th>
<th>274-dim</th>
<th>324-dim</th>
<th>average</th>
</tr>
</thead>
<tbody>
<tr>
<td>PQ method (full embedding)</td>
<td>2000</td>
<td>70</td>
<td>73</td>
<td>58</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>6000</td>
<td>76</td>
<td>77</td>
<td>84</td>
<td>79</td>
</tr>
<tr>
<td></td>
<td>10 000</td>
<td>80</td>
<td>79</td>
<td>91</td>
<td>83</td>
</tr>
<tr>
<td>Average</td>
<td>6000</td>
<td>75</td>
<td>76</td>
<td>77</td>
<td>76</td>
</tr>
<tr>
<td>MB method (full embedding)</td>
<td>2000</td>
<td>71</td>
<td>66</td>
<td>90</td>
<td>76</td>
</tr>
<tr>
<td></td>
<td>6000</td>
<td>77</td>
<td>75</td>
<td>95</td>
<td>82</td>
</tr>
<tr>
<td></td>
<td>10 000</td>
<td>88</td>
<td>81</td>
<td>99</td>
<td>89</td>
</tr>
<tr>
<td>Average</td>
<td>6000</td>
<td>79</td>
<td>74</td>
<td>95</td>
<td>82</td>
</tr>
<tr>
<td>Contourlet-based method (full embedding)</td>
<td>2000</td>
<td>56</td>
<td>69</td>
<td>68</td>
<td>64</td>
</tr>
<tr>
<td></td>
<td>6000</td>
<td>59</td>
<td>74</td>
<td>95</td>
<td>76</td>
</tr>
<tr>
<td></td>
<td>10 000</td>
<td>62</td>
<td>75</td>
<td>98</td>
<td>78</td>
</tr>
<tr>
<td>Average</td>
<td>6000</td>
<td>59</td>
<td>73</td>
<td>87</td>
<td>73</td>
</tr>
<tr>
<td>YASS method (full embedding)</td>
<td>2000</td>
<td>57</td>
<td>57</td>
<td>59</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td>6000</td>
<td>62</td>
<td>63</td>
<td>60</td>
<td>62</td>
</tr>
<tr>
<td></td>
<td>10 000</td>
<td>61</td>
<td>69</td>
<td>65</td>
<td>65</td>
</tr>
<tr>
<td>Average</td>
<td>6000</td>
<td>60</td>
<td>63</td>
<td>61</td>
<td>61</td>
</tr>
<tr>
<td>Scheme no. 1 (partial embedding) (MB)</td>
<td>2000</td>
<td>65</td>
<td>70</td>
<td>55</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td>6000</td>
<td>71</td>
<td>75</td>
<td>79</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td>10 000</td>
<td>75</td>
<td>76</td>
<td>86</td>
<td>79</td>
</tr>
<tr>
<td>Average</td>
<td>6000</td>
<td>70</td>
<td>74</td>
<td>73</td>
<td>72</td>
</tr>
<tr>
<td>Scheme no. 2 (MB, PQ)</td>
<td>2000</td>
<td>62</td>
<td>66</td>
<td>55</td>
<td>61</td>
</tr>
<tr>
<td></td>
<td>6000</td>
<td>66</td>
<td>71</td>
<td>73</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>10 000</td>
<td>69</td>
<td>75</td>
<td>81</td>
<td>75</td>
</tr>
<tr>
<td>Average</td>
<td>6000</td>
<td>66</td>
<td>71</td>
<td>70</td>
<td>69</td>
</tr>
<tr>
<td>Scheme no. 3 (MB, PQ, Contourlet-based)</td>
<td>2000</td>
<td>62</td>
<td>61</td>
<td>58</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>6000</td>
<td>67</td>
<td>70</td>
<td>68</td>
<td>68</td>
</tr>
<tr>
<td></td>
<td>10 000</td>
<td>73</td>
<td>76</td>
<td>74</td>
<td>74</td>
</tr>
<tr>
<td>Average</td>
<td>6000</td>
<td>67</td>
<td>69</td>
<td>67</td>
<td>68</td>
</tr>
<tr>
<td>Scheme no. 4 (MB, PQ, YASS)</td>
<td>2000</td>
<td>56</td>
<td>57</td>
<td>58</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td>6000</td>
<td>59</td>
<td>64</td>
<td>60</td>
<td>61</td>
</tr>
<tr>
<td></td>
<td>10 000</td>
<td>62</td>
<td>68</td>
<td>63</td>
<td>64</td>
</tr>
<tr>
<td>Average</td>
<td>6000</td>
<td>59</td>
<td>63</td>
<td>60</td>
<td>61</td>
</tr>
<tr>
<td>Scheme no. 5 (PQ, YASS, Contourlet-based)</td>
<td>2000</td>
<td>55</td>
<td>54</td>
<td>55</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>6000</td>
<td>57</td>
<td>60</td>
<td>59</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td>10 000</td>
<td>59</td>
<td>64</td>
<td>63</td>
<td>62</td>
</tr>
<tr>
<td>Average</td>
<td>6000</td>
<td>57</td>
<td>59</td>
<td>59</td>
<td>58</td>
</tr>
</tbody>
</table>

Images indicate that the size of securely embedded secret data with the proposed hybrid steganography approach is more than the size of secret data embedded with traditional steganography techniques.

Figure 3 illustrates the comparison of WBS, 274-dim, and 324-dim steganalyzers’ average accuracy in detection of different embedding schemes of HYSA. As the diagram shows, scheme no. 1, which is the partial embedding of MB method, produces stego images (partial stego) that have higher undetectability than stego images constructed by full embedding (full stego) of MB method. In addition, stego images produced by schemes no. 2–5, which are combinations of different steganography methods, have more

Figure 3. Comparison of WBS, 274-dim, and 324-dim steganalyzers’ average accuracy in detection of different embedding schemes.
security than stego images produced by full embedding of MB, PQ, Contourlet-based, or YASS methods.

Figure 4 shows the features’ values of clean, partial stego, and full stego images, respectively, for a randomly selected image from the database. In both partial and full stego images, 10,000 bits are hidden. As we see in Figure 4, the values of steganalyzer features in clean and partial stego images are closer to each other compared to the feature values of full stego image. This means that the statistical features of images in partial embedding are altered not as much of full embedding. The differences between values of features are different in various images. Some features of an image may vary a lot in clean, partial, and full embedding but in other images the value of some other features may change. In Figure 4(b) and (c), which show 324-dim and 274-dim steganalyzers, for the features that are not shown in the figure, the values in all three images (clean, partial stego, and full stego) are nearly equal or the values in partial and full stego images are similar.

Figure 5 shows the diagrams for Baboon clean image, partial embedding in Baboon image (in quarter of the image) with 10,000 bits, and full embedding in Baboon image...
Figure 5. Comparison of features’ values of (a) WBS, (b) 324-dim, (c) 274-dim steganalyzers for clean, partial, and full embedding in Baboon image using MB steganography method.

(in whole of the cover image) with 10 000 bits. Each diagram compares the value of features of WBS, 324-dim, and 274-dim steganalyzers for clean, partial stego, and full stego images. The range of WBS feature values is from $-2$ to 4040 and the range of 274-dim feature values is from $-1726$ to 14 131. Therefore, for convenience we show the features’ values of these steganalyzers in percentage. The vertical axis in Figure 5(a) and (c) have been normalized to $[-100, +100]$, due to the wide range of feature values in WBS and 274-dim steganalyzers.
Generally, other typical images, such as Lena, Jet, etc., have similar diagram behaviors with minor differences. As Figure 5 illustrates, in almost all images, the values of features extracted from partial stego images is close to those of clean (cover) images and the values of features extracted from full stego images are farther from clean images.

Figure 6 demonstrates the outcome of HYSA compared to single method embedding. As the results show, in most cases, the difference between the features of clean image and features of hybrid embedded stego image is lower than the difference of clean image features and single embedded stego image features. In all three diagrams, first, the feature values are normalized and then the differences are calculated.

In the steganography arena, steganography–steganalysis is a ‘cat and mouse’ game: knowing the steganography method, a technique can be devised to detect it; and knowing the detection method, the steganography algorithm can be further modified to mislead the detection procedure [25]. Therefore, a steganalysis method can be proposed later to detect HYSA such as the detection of MB, PQ, and YASS algorithms. In the future, we will try to oblige HYSA to be undetectable against the future steganalyzers.

Figure 6. The difference of features’ values of clean image and single method embedding stego image (solid line) and difference of features’ values of clean image and hybrid embedding stego image (dash line) in (a) WBS, (b) 324-dim, (c) 274-dim steganalyzers.
5. CONCLUSION

To have a protected steganography method from efficient and well-known steganalyzers, one may propose an embedding method that minimizes the statistical artifacts or organize an approach that confuses the steganalyzers estimations. Considering the operation of steganalyzers, in this paper we proposed HYSA. In this approach, initially the steganographer selects one or more predefined steganography methods randomly and then he chooses one or more random regions in the cover image. In next step, the secret data are embedded in the chosen regions with selected steganography methods. Consequently, the risk of detection will be decreased because of two reasons. First, the statistical features are changed only in some regions and the value of features that are calculated based on the overall properties of stego images are changed slighter and are not large enough to be detectable by steganalyzers. Second, different steganography methods apply dissimilar embedding techniques and each method affects on various statistical features of images. Therefore, employing HYSA may cause irregular and hard to detect statistical artifacts compared to utilizing a single embedding method. According to the obtained results, for the same detection accuracy, scheme no. 5 provides more undetectable stego images than employing PQ, MB, or Contourlet-based steganography method. Additionally, partial embedding (scheme no. 1) produces more secure stego images compared to full embedding (classical employing) by steganography methods. In the future, we will try to strength HYSA to be undetectable against the upcoming steganalysis algorithms.

REFERENCES

4. Hetzl S, Mutzel P. A graph theoretic approach to steganography. In Proceedings of 9th International Con-
Hybrid steganographic approach


