ISSN No. 0976-5697

Volume 6, No. 8, Nov-Dec 2015

International Journal of Advanced Research in Computer Science

REVIEW ARTICLE

Available Online at www.ijarcs.info

Study of Image Steganalysis Techniques

Archana O. Vyas
Deptt. of Electronics and Tele-Communication Engg.,
G. H. Raisoni College of Engineering and Management,
Amravati, India

Dr. Sanjay V. Dudul
Deptt. of Applied Electronics,
Sant Gadge Baba Amravati University,
Amravati, India

Abstract: With the emergence of steganography, the counter technology, namely steganalysis has also emerged. Steganalysis is used to detect or extract the hidden message from the carriers. It is a set of techniques: visual or statistical by which it is possible to check for the existence of steganography content in cover object. Steganalysis could be passive or active. Passive steganalysis simply aims to identify the presence or absence of secret message whereas Active steganalysis attempts to estimate the message length, secret key, message bits, etc. Current steganalysis techniques emphasize on the design of the classifier based on the training set of cover objects and stego objects obtained from a variety of different embedding algorithms. The inherent features of natural image get violated when an image undergoes some embedding process; classification is done on the basis of the some of these features. The extraction of sensitive features and design of good classifier are the principal tasks for steganalysis.

Keywords: Steganalysis; Classifier; Feature extraction; RBFNN; LGEM; GEFR; LSB; GA; RS; POV

I. INTRODUCTION

Although steganography may provide a safe tool for communication to government and business, it will suffer serious consequences if it is used by the terrorists or criminals. In contrast, steganalysis was proposed to determine whether there are secret messages embedded in image or video [1].

Steganalysis involves two major types of analysis: Visual analysis and statistical analysis. Visual analysis deals with detection of secret message with naked eye or with the help of computer in which bit planes are analyzed separately for any unusual change in the appearance for the presence of secret message. Statistical analysis deals with detection of any change in statistical properties of stego object caused by steganographic algorithm. Steganalysis can be divided into two major types: Universal steganalysis and specific type of steganalysis techniques. Universal steganalysis techniques can detect secret message in stego objects embedded by a range of steganographic algorithms and specific stegalalysis techniques, which are more sophisticated techniques and work corresponding to a particular steganographic algorithm only [2].

Universal steganalysis is also referred as blind steganalysis because these are independent of any specific embedding technique and are used to alleviate the deficiency of targeted analyzers by removing their dependency on the behavior of individual embedding techniques. To achieve this, a set of distinguishing statistics that are sensitive to a wide variety of embedding operations are determined and collected. These statistics, computed from both cover and stego images are employed to train a classifier, which is subsequently used to distinguish between cover and stego images [3].

Blind steganalysis has two important components; these are feature extraction and feature classification. In feature extraction, a set of distinguishing statistics are obtained from a data set of images. There is no well defined approach for obtaining these statistics, but often they are proposed by observing general image features that show strong variation under embedding. The second component, feature classification, operates in two modes. First, the obtained distinguishing statistics from both stego and cover images are used to train a classifier. Second, the trained classifier is used to classify an input image as either being a clean image or carrying a hidden message [3].

Specific type of steganalysis techniques or embedding specific steganalysis is also termed as Targeted steganalysis and are designed for a particular steanographic algorithm. This technique is more robust, since it has good detection accuracy for that specific technique when they were used against the particular steganographic technique [3].

Steganalysis technique can also be classified as whether they use instances to build classifiers or not. The steganalysis approaches can be categorized into two types, one uses instances of training images, the other does not use instances, instead, a parametric model is developed and its statistics are computed for steganalysis detection. Non-instance based techniques use the statistics of the image by an implicated parametric model, and classification is based on rules [4].

Computational intelligence approach for steganalysis:

Data hiding process distorts the bit plane of the cover signal. If the host signal is an image, then the difference between images before and after the embedding process may give some information about the hidden data. That is if the bit planes of the two images are compared, the existence of the hidden data can be discovered. Rather than making an exact comparison between clean and altered images, ANN can be used to learn the bit pattern of images that makes the trained ANN capable of classifying the images. An essential issue in training neural networks is to choose training inputs in such a way that the training error is minimized. This selection needs to take proper care of the features of images [4].

The detection of hidden information detection is regarded as a classification process, in this approach. The

inputs are the images and the outputs are the class labels. Feature extraction is often used in blind steganalysis which aims to build a universal classifier to detect all steganography without knowing the actual methods used. Classifiers are used to wrap the extracted feature. The classifier is designed using the training set, some classified instances are applied, and then the classifier is used to classify unknown test set. A variety of classification methods can be used and different classification algorithms can be designed for better performance of the steganalyzer [5]. For steganalysis, the spatial domain & transform domain features extracted from the stego image could be applied to the neural network classifier in order to classify the original and the stego image. Suitable learning machine may be applied as a classifier for the desired steganalysis process. The general computational intelligence based approach for steganalysis can be illustrated as in figure 1.

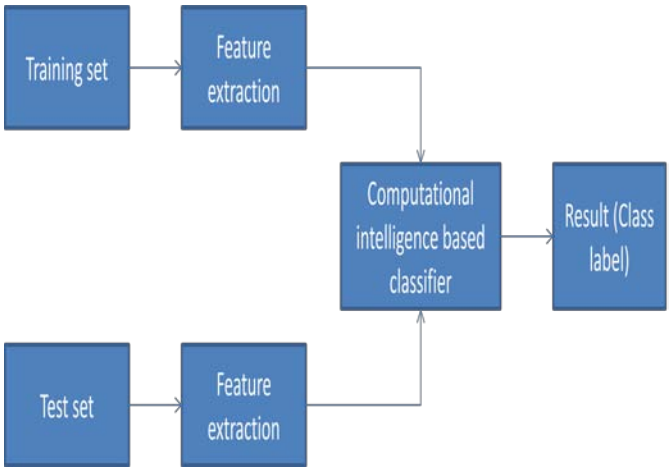

Figure1 – Computational Intelligence based LSB steganographic detection[4]

The rest of this paper is organized as follows:
Section II highlights various image steganalysis methods. Section III explains various performance parameters used to be evaluated for steganalysis. Section IV elaborates the scope of research noticed so far in the reported research work in the field of steganalysis. Finally, section V describes the concluding part.

II. STEGANALYSIS METHODS

The performance of the steganalysis system is mainly determined by the method of feature extraction and the architecture selection of the classifier. Selecting the proper classifier with appropriate parameters will improve the detection accuracy and generalization capability of the system [6].

A. Steganographic detection by Radial basis function neural network minimizing the localized generalization error model:

Radial Basis Function Neural Network (RBFNN) optimized by the Localized Generalization Error Model (L-GEM) is being used by various researchers for steganography detection. In this method discrete cosine transform (DCT) features along with the Markov features were used as inputs of neural networks for detection. To enhance the generalization capability of the RBFNN, the architecture of the RBFNN is selected by minimizing the L-GEM. The Discrete Cosine Transform (DCT) features and Markov features are extracted from JPEG images as the input features for RBFNN describing an image. To enhance the generalization capability of steganography detection, architecture selection of RBFNN was done using the L-GEM. RBFNN could be trained quickly for large dataset and it is not easily trapped by the local minima. The L-GEM estimates the generalization error of a RBFNN for unseen samples, which are similar to the training samples. In practical applications, one may not expect a classifier to correctly classify unseen samples, which are very different from the training samples. Therefore, selecting classifier architectures based on the generalization error of all possible samples in the entire input space may be counter-productive [6].

DCT features focus on capturing the statistic of DCT coefficients and the inter-block dependencies among DCT coefficients. On the other hand, Markov features capture the intra-block dependencies. Both DCT features and Markov features have their own superiority for dealing with different types of steganographic algorithm. Merging DCT features and Markov features would have improved the detection accuracy and the robustness of the system [6].

B. Steganographic detection using Bayesian neural network and genetic algorithm:

A feature selection and transformation approach for universal steganalysis based on Genetic Algorithm (GA) and higher order statistics have been reported in [7] [8].

They have chosen three types of typical statistics as candidate features and twelve kinds of basic functions as candidate transformations. The GA is utilized to select a subset of candidate features, a subset of candidate transformations and coefficients of the Logistic Regression Model for blind image steganalysis. The Logistic Regression Model is then used as the classifier. They have observed that the GA based approach increases the blind detection accuracy and also provides a good generality by identifying an untrained stego-algorithm.

c. LSB Steganographic detection using non instance based techniques:

Most of the conventional LSB steganalysis methods fall in this category. Three of them are illustrated below.

Chi-square steganalysis:

LSB utilizes the lowest bit plane for embedding, and substitutes every suitable bit with a message bit. If the embedded message bit is different from the original one, then a bit flip must have been performed [5]. Let the pixel value be j, $j \in [0.255]$. If $j = 2i$ then after flipping j will be $2i+1$, and if $j = 2i + 1$ then after flipping j will be $2i$. According to concept of POV (Pairs of Values), which combines two pixel values $2i$ and $2i + 1$ together as a pair, and the two values in the pair only differ in the lowest bit. It can be easily noticed that if a pixel value is in a pair, then after embedding the value is still in the same pair. Let the frequencies of pixel value be $2i$, $2i+1$ be h_{2i}, h_{2i+1}. Usually, messages are compressed or encrypted before embedding,

so the message bits can be treated as random, and subsequently the frequencies of 0 and 1 would be nearly 1/2. So that is to say, the frequencies of the two pixel values in the same POV would be nearly the same after embedding. In order to determine whether there exists significant difference between h_{2i} and h_{2i+1}, χ^2 test could be used. It is needed to calculate χ^2 statistics as (after embedding the expectation of h_{2i} is, the probability of image being embedded can be calculated using the density function of χ^2 distribution as $p(k$ is the total number of all possible i in equation 1): If p is close to 1, the image is classified as an embedded stego. χ^2 steganalysis is used to estimate the hidden message length, but it can also be used to detect the existence of hidden messages. The piecewise χ^2 coefficients and a threshold could be chosen, if there are coefficients which are higher than the threshold, then the image is classified as embedded one [5].

Regular Singular (RS) steganalysis:

RS method exploits the spatial correlations in stego-images. Let us assume that we have $M \times N$ images, and let P stands for all pixel values. We define discrimination function f on group $G = (x_1,\ldots\ldots, x_n)$ as

$$f(x_1,\ldots. x_n) = \sum_{i=1}^{n-1} |xi - xi + 1|,$$

(i = 1, 2, …..., n).The discriminate function f can be regarded as a smoothness measure of G. There are three types of flipping functions defined. $F1$ is defined as↔ 0 1; 2↔3,……,254↔255; F_{-1} as -1↔0, 1↔2,……,255↔256; F_0 is defined as identity permutation $F_0(x) = x$. So LSB embedding can be expressed as follows [5]:

When the embed bit and corresponding cover bit are the same, F_0 is applied, else $F1$ is applied. LSB embedding will cause an increase of the f value. Every group is applying flipping functions using mask M and as a result flipped group is $F(G) = (F_{M(1)}(x1), F_{M(2)}(x2),\ldots\ldots, F_{M(n)}(xn))$ where M (i) € {-1,0,1}. Three types of pixel groups R, S, U are defined as Regular groups: $G€R \leftrightarrow f(F(G)) > f(G)$, Singular groups: $G€S \leftrightarrow f(F(G)) < f(G)$ and Unusable groups: $G€U \leftrightarrow f(F(G)) = f(G)$. The relative number of 'R' groups of non-negative mask M€{0,1} are denoted as R_M, S groups as S_M. Similarly the relative number of R groups of non- positive mask $-M$€{-1,0} can be denoted as R_{-M}, S group as S_{-M}. The LSB embedding enforces the difference between R_M and S_M to zero as embedding length increases [5].

The Sample Pair analysis (SPA):

The SPA method traces the multisets of sample pairs before and after LSB embedding, and uses relationships between multisets to solve the length of embedded messages.

Images could be represented by successive samples s_1, s_2, …..s_N (N is the total number of the sample). A sample pair is a tuple $(s_i; s_j), 1 \le i, j \le N$, and let P be the set of all sample pairs drawn from the image. D_n is defined to be the submultiset of P which contains sample pairs like $(u, u+n)$ or $(u+n, u)$ where $0 \le n \le 2^b - 1$(b is the total bit number). For $0 \le m \le 2^{b-1} - 1$, C_m is defined as the sample pairs whose values differ by m in the first (b -1) bits (i.e. by right shifting one bit and then difference is measured). $X_{2m+1} = D_{2m+1} \cap C_{m+1}$ and $Y_{2m+1} = D_{2m+1} \cap C_m$, is defined for $0 \le m \le 2^{b-1} - 2$and $X_2{}^b{}_{-1} = 0$, $Y_2{}^b{}_{-1} = D_2{}^b{}_{-1}$. So if the sample pairs of P are uniformly scattered, then for m, $0 \le m \le 2^b - 2$ it turns out that,

$$E|X_{2m+1}| = E|Y_{2m+1}|$$

This is the key observation of SPA method [5].

Table1: Comparison between Conventional LSB steganalysis for Detection[5]

Method	**Principle**	**Feature**	**Classification for Detection**	**Using Machine learning**
χ^2	Split images into segments and use χ^2 test for every segment.	χ^2 Coefficients.	Threshold based.	No
RS	Use mask and flip to identify R,S,U groups and draw RS diagram to estimate.	The frequency of R,S,U groups in different masks.	Threshold based.	No
SPA	Calculate the frequency of sample pairs and solve equations to estimate.	The frequency of every specific sample pair.	Threshold based.	No

d. LSB Steganographic detection using Gradient Energy Flipping Rate (GEFR):

The spatial Least Significant Beat (LSB) Steganography results in the alteration of the smooth characteristics between adjoining pixels of the raw image. The relation between the length of embedded message and the gradient energy is theoretically analyzed, and then a steganalysis and detection method, named Gradient Energy-Flipping Rate detection (CEFR) was proposed in [9]. Through the analysis of the variation of the gradient energy, which results from the LSB Steganography in colour and grayscale image, the secret message embedded in the target image could be detected, and the length of the embedded message could be estimated. The method was found to have detection rate 0.0l bit per pixel [9].

e. Steganalysis of Pixel Value Differencing (PVD) steganographic method.

PVD through a zigzag scan of the image generates a vector which is called the "image vector". The difference between every pair of elements of this vector would produce yet another vector which is termed as the "substitute vector". In PVD steganography, the embedding is performed on this vector. An image could be built out of a substitute vector and called the "substitute image". PVD steganalysis method is based on the construction of this image. The embedding is performed only in permissible elements of the substitute vector. Each of those elements would contain a variable number of embedded data. Later this vector is combined with the cover image to produce the stego image [10].

For any image, a substitute vector could be produced and substitute image out of it could be formed. In construction of the substitute image, only elements of the substitute vector are involved that are eligible for embedding. The

histogram of the substitute image should contain bin pairing if the image under analysis contains embedding through PVD algorithm. Hence, PVD steganalysis is based on histogram changes. Suppose p_i denotes the i th pixel of the image vector and p'_j is the j th element of the substitute vector. It is been observed that the histogram of the differential for regular images with no embedding has a Gaussian like or similar distribution [10]. With increase in the amount of embedding, the histogram involves pairing in its adjacent bins. The pairing effect is well distinguishable for the image with embedding.

III. PERFORMANCE PARAMETERS

The effectiveness of a steganalysis technique can be evaluated based on the following performance parameters.

- *Detection Accuracy:*
- *Detection Rate:*
- *True Positive Rate (TPR):* It is defined as the ratio of the number of correctly classified images out of the overall test images. TPR should be as high as possible.
- *False Positive Rate (FPR):* It represents the ratio of the wrongly classified the plain images as stego ones. FPR should be as low as possible.
- *Classification Rate:* Classification rate is defined as the average of positive detection (PD) and negative detection (ND), given by

(PD + ND) / 2

Where, Positive Detection (PD) is classifying the stego images correctly and Negative Detection (ND) is classifying the non stego images correctly [11].

IV. SCOPE OF FURTHER RESEARCH

The literature survey provides details of how algorithms have evolved over time with respect to the nature of stego objects. In addition, preceding sections also shed light on some characteristics that are extremely necessary for good steganalysis system. Incorporating all of these features into a single system is itself a matter of significant research. However, in the light of information gathered hitherto, some of the possibilities of future scope in the field of digital image steganalysis are listed below.

- Finding out more effective features to detect the existence of secret messages embedded with most kinds of data embedding schemes [11].
- Identifying the type of steganographic algorithm utilized to generate the stego image and locating the image regions exploited to hide secret messages [11].
- It could be tried to improve the embedding length estimation precision especially when the embedding length is relatively short [12]
- More effective learning techniques like cost sensitive learning and class-imbalance learning could be incorporated [5].

V. CONCLUSION

Image steganography techniques have been applied to several fields like copyright protection, fingerprinting, digital rights measurement (DRM), etc. But the misuse of this technology is a threat to national security as well as social security. So, the current research focuses on analysis of the stego images thus called steganalysis. This paper overviews the most established algorithms for image steganalysis based on different feature extracted and different classifier designs. Based on the information gathered through the analysis, some important characteristics of a good steganalysis system have been figured out and future possibilities of research in the area of image steganalysis have also been pointed out.

VI. REFERENCES

[1.] Zhi-min he, Wing w.y. ng, Patrick P.K.chan and Daniel s. yeung, **"**feature selection for blind steganalysis using localized generalization error model", IEEE, Proceedings of the ninth International Conference on Machine Learning and Cybernetics, Qingdao, 11-14 July 2010.

[2.] SumeetKaur, SavinaBansal and R. K. Bansal,"Steganography and Classification of Image Steganography Techniques", 2014, IEEE International Conference On Computing For Sustainable Global Develpoement (INDIACom), 978-93-80544-12-0/14.

[3.] Deepa D. Shankar, T. Gireeshkumar, K. Praveen, R. Jithin, and Ashji S. Raj, "Block Dependency Feature Based Classification Scheme for Uncalibrated Image Steganalysis", 2012: ICDEM 2010, LNCS 6411, pp. 189–195, 2012. © Springer-Verlag Berlin Heidelberg 2012

[4.] Hossein Malekmohamadi and Shahrokh Ghaemmaghami, "Reduced Complexity Enhancement Of Steganalysis Of LSB-Matching Image Steganography", IEEE transaction on computer system and application, AICCSA, 2009, 978-1-4244-3806.

[5.] Shen Ge, Yang Gao and Ruili Wang, "Least Significant Bit Steganography Detection with Machine Learning Techniques", 2007 ACM SIGKDD Workshop on Domain Driven Data Mining (DDDM2007), August 12, 2007, San Jose, California, USA.

[6.] Zhi-min he, Wing w.y. ng, Watrickp.k. chan, Waniel s. yeung, "Steganography Detection Using Localized Generalization Error Model", IEEE transaction on systems man and cybernetics, 978-1-4244-6588.

[7.] Xiao Yi Yu, Aiming Wang, "Steganalysis Based on Bayesion Network and Genetic Algorithm", Sponsored by Program for Science and Technology Innovation Talents in Universities of Henan Province, and Tian Jin Natural Science Fundation, 978-1-4244-4131-0/09/$25.00 ©2009 IEEE.

[8.] Xiao Yi Yu, Aiming Wang, "An Investigation of Genetic Algorithm on Steganalysis Techniques", IEEE, 2009 Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing.

[9.] Wen-Nung Lie and Guo-Shiang Lin, "A Feature-Based Classification Technique for Blind Image Steganalysis", IEEE Transactions on Multimedia, vol. 7, no. 6, December 2005.

[10.] Vajiheh Sabeti, Shadrokh Samavi, Mojtaba Mahdavi and Shahram Shirani, "Steganalysis of Pixel-Value Differencing Steganographic Method", IEEE transaction on communications computers and signal processing, 2007,1-4244-1190.

[11.] Wen-Nung Lie and Guo-Shiang Lin, "A Feature-Based Classification Technique for Blind Image Steganalysis", IEEE Transactions On Multimedia, vol. 7, no. 6, December 2005.

[12.] Li Zlii, Sui Ai Fen and Yang Yi Xian, "A LSB Steganography Detection Algorithm", The 14th IEEE, 2003, International Symposium on Persona1,lndoor and Mobile Radio Communication Proceedings.

Ms. Archana O. Vyas was born in Indore, Madhyapradesh in 1980. She received the B.E. Degree in Electronics and Tele-communication from S.G.B. Amravati University, Amravati in 2009 and completed her M.Tech. in Electronic Systems and Communication from Government college of Engineering Amravati in 2011. Currently she is working as Assistant Professor in Electronics and Tele communication Engg. Department at G. H. Raisoni College of Engineering & Management, Amravati. She is pursuing PhD. degree in Electronics Engineering from Sant Gadge Baba Amravati University, Amravati, India. Her interest of research is image steganography and steganalysis using computational intelligence approach.

Sanjay V. Dudul was born on 28th August 1964 at Amravati, Maharashtra. He was conferred on Ph.D. in Electronics Engineering on topic of Computational Intelligence by Sant Gadge Baba Amravati University, Amravati in 2003. He obtained B.E. degree in Electronics Engineering in 1986 from Nagpur University. He received M.E. degree in Electronics Engineering with specialization in Computer Applications in 1989. Currently, he is the Professor and Head of the Department of Applied Electronics at SantGadage Baba Amravati University, Amravati, India. He has published 81 research articles in peer reviewed and refereed International journals and 12 research articles in refereed National Journals. His numerous publications have been listed in SCOPUS. Hitherto, 10 students have been awarded Ph.D. degree in Electronics Engineering under his guidance. He has filed 5 Indian patents. He is a member of editorial boards of many International Journals. His research area includes Image Processing, AI in Pattern Recognition, Prediction and Regression, Machine Intelligence: A Soft-Computing Approach, Advance digital signal processing, Pattern Recognition, Analysis and Machine Intelligence, Computational Intelligence, Bio-medical Signal Processing, Intelligent/Smart Sensors and Learning Systems.