International Journal of Advanced Computer Research (ISSN (print): 2249-7277 ISSN (online): 2277-7970)
Volume-3 Number-3 Issue-12 September-2013

Effective Steganalytic Algorithm based on Residual Value of Pixels in an Image

T. J. Benedict Jose[1], P. Eswaran[2]

Abstract

Steganalysis is the art and science of detecting the presence of secret information in the cover object. In this work, we propose a steganalytic algorithm, which differentiates the normal image from the stego image. To achieve this, we train a set of images by applying a model that finds all the dependencies among pixels. This model, then calculates the residual value along 0, 45, 90 and 135 degrees. The calculated residual value is then converted into a single digit for memory reduction. These residual values are considered as features. In testing phase, the residual value is calculated for all the pixels and the training feature value is loaded. SVM classifier is employed to differentiate between the training and testing feature. On the basis of minimal distance value, it is decided whether or not the image is stego.

Keywords

Steganalysis, SVM Classifier, Stego image, Residual Value.

1. Introduction

Steganalysis is an art and science, which discriminates between the stego and the cover objects. This analysis has to be done without any prior knowledge of the secret key being used to embed and also the embedding algorithm.

Mostly, images are used as cover objects because of its wide usage. Using image as carrier, which is so called 'cover image', some secret message is embedded by using an embedding algorithm in cover image, in order to arrive at a stego image. The process of embedding a secret message into a cover medium is steganography and the science of identifying the presence of secret message is steganalysis.

T.J Benedict Jose, Department of Computer Science, Manonmaniam Sundaranar University, Tirunelveli, India.
P Eswaran, Department of Computer Science, Alagappa University, Karaikudi, India.

In this work, the steganalysis is done in an image, which has undergone LSB (Least Significant Bit) Steganography. In this technique, the lowest bit plane of a bitmap image is used to convey the secret data, which the human eye cannot detect easily.

Data Embedding
The scheme as shown in Fig 1, initializes some parameters, and these parameters are used for data preprocessing, region selection etc., after this, the capacity of the selected region is figured out [16].

If the regions are capable then the given secret message is embedded. A stego image is obtained after performing some kind of post processing.

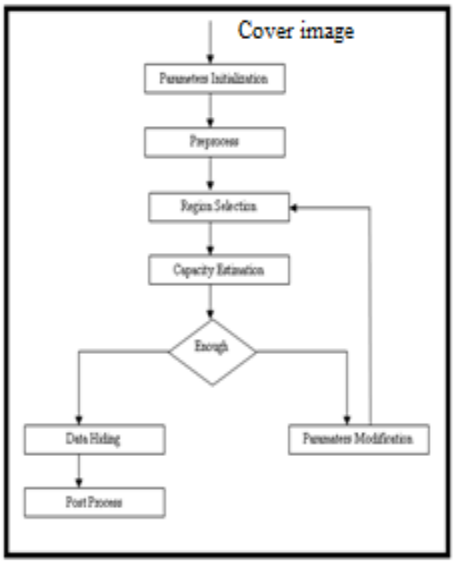

Fig 1: Data Embedding

Data Extraction
Fig 2 illustrates how the data is extracted. To extract data, the scheme as given in Fig 2, first extracts the side information from the stego image. Based on the side information, it then does some preprocessing and identifies the regions that have been used for data hiding. Finally, it obtains the secret message M according to the corresponding extraction algorithm.

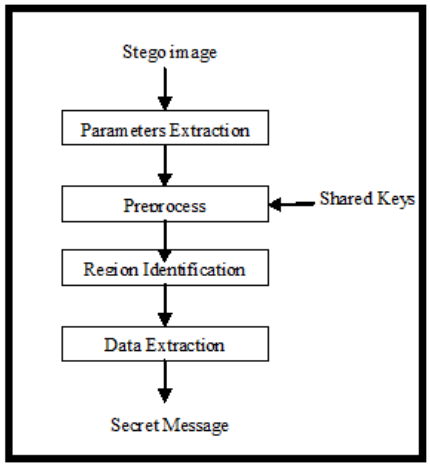

Fig 2: Data Extraction

By using the above mentioned steganographic algorithm, the secret message is embedded and extracted successfully. Important merits of this scheme are detection of secret message is slightly complex, security performance is better.

This scheme provides more embedding capability. Thus, this steganographic algorithm is employed in our work and our objective is to detect the presence of secret message and this is done by our proposed steganalytic algorithm. SVM Classifier is used to differentiate between the normal and stego images.

2. Related Work

Fridrich[1] have broken the F5, which is a steganographic algorithm, by estimating the value of 'β', that can consequently be turned into an estimate of the secret embedded message.

After getting the baseline histograms, the modified non-zero non DC coefficients are determined, in order to arrive at 'β', which minimizes the least square error between the stego image histogram and the histograms obtained by embedding a message that paves way to extract 'β, modifications.

Selection of cover images is also given importance in the work done by Fridrich. Selection of cover images has its impact over the stego system and security. Images with low number of colors, computer art are needed to be excluded, since they cannot serve as better cover images.

Aura[7] suggests that grayscale images are the best choice for cover images. Fridrich[8] stated that decompressed JPEG images should be avoided as cover medium for spatial steganography, such as the LSB embedding or so. Westfeld[9] presented a method that is based on statistical analysis of Pairs of Values, which are exchanged when the message is embedded.

This method is known as $\chi 2$ attack which makes use of correlations between input and output measured by the $\chi 2$ test. The $\chi 2$ -attack was originally proposed by Vaudenay as an attack on the Data Encryption Standard (DES) and Handschuh et al. This can be applied to many embedding paradigms rather than the LSB embedding. Farid[10] has implemented a detection scheme, which is applicable to all steganographic schemes.

This is made possible by providing proper training of original and stego images to databases. Farid made use of optimal linear predictor to have wavelet coefficients, based on which the first four moment of the distribution of prediction error is calculated.

Then a threshold is found out by using Fisher's linear discriminant statistical clustering, to identify between the stego and the normal images.

Westfeld[2] presented a new steganographic algorithm named F5, which emphasized that the embedding density should remain the same throughout. This resulted in a high steganographic capacity with high efficiency, since more number of bits can be embedded per change. Huang[3] presented a new channel rule for JPEG steganography, which reduced the detection capability of several JPEG steganalyzers.

The efficiency of this channel is tested with four different JPEG steganalyzer. Fridrich[4] investigated the statistical detectability of several steganographic methods, by determining maximal relative payload, at which the methods would be undetectable statistically.

Fridrich's work is mainly meant to detect between the normal and stego images by exploiting SVM, in which a model is obtained by training a dataset, and then the obtained model is used over the testing dataset, to predict information [11].

3. Proposed Work

In our proposed work, residual value is calculated along 0, 45, 90 and 135 degrees, for improved accuracy in detecting the stego image. All the steps involved are listed below.

Training Phase

1. In this phase, the name of the folder that contains training images is given as input.
2. Then, the below steps are applied for all images in the training folder

a. Cover the image into matrix
b. Apply the rich model for all pixels in the image for a large number of different types of dependencies among neighbouring pixels to give the model
c. The rich model calculates the residual along 0,45,90,135 degree directions
d. The rich model is shown in the below figure

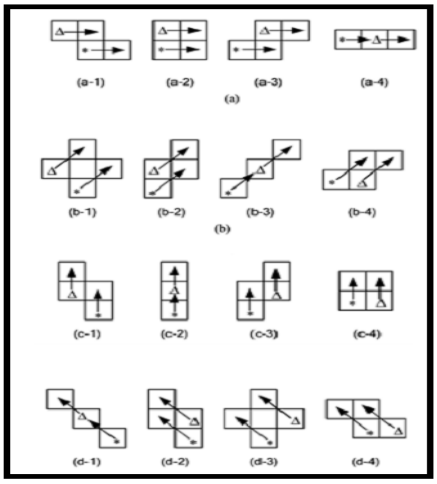

Fig 3: Model (a) represents the 0 degree direction,(b) model represents the 45 degree directions, (c)model represents 90degree directions (d) model represents the 135 degree directions.

From each of the above model, the residual value is calculated. This residual value is converted into the single digit value for memory space reduction. The residual value is the difference between the centre pixel value and its neighbouring pixel value.

3. These residual values are considered as the features
4. These residual values are saved for testing phase.

Testing Phase

1. Get the testing image which is to be analysed to check whether or not it is a stego image.
2. Calculate the residual value for all pixels as well as for the the training phase
3. Load the training feature value
4. Apply the Support Vector Machine (SVM) classifier to find the difference between the training and testing features.
5. Decide that the image is a stego or normal image based on the minimum distance value.

SVM Classifier

We use SVM classifier for finding the difference between the training and the testing features. Some of the reasons for why SVM Classifier is employed are as follows.

1. By introducing the kernel, SVMs gain flexibility in the choice of the form of the threshold separating solvent from insolvent companies, which needs not be linear and even needs not have the same func- tional form for all data, since its function is non-parametric and operates locally. As a consequence they can work with financial ratios, which show a non-monotone relation to the score and to the probability of default, or which are non-linearly dependent, and this without needing any specific work on each non-monotone variable.
2. Since the kernel implicitly contains a non-linear transformation, no assumptions about the functional form of the transformation, which makes data linearly separable, is necessary. The transformation oc- curs implicitly on a robust theoretical basis and human expertise judgement beforehand is not needed.
3. SVMs provide a good out-of-sample generalization, if the parameters C and r (in the case of a Gaussian kernel) are appropriately chosen. This means that, by choosing an appropriate generalization grade, SVMs can be robust, even when the training sample has some bias.
4. SVMs deliver a unique solution, since the optimality problem is convex. This is an advantage compared to Neural Networks, which have multiple solutions associated with local minima and for this reason may not be robust over different samples.

5. With the choice of an appropriate kernel, such as the Gaussian kernel, one can put more stress on the similarity between companies, because the more similar the financial structure of two companies is, the higher is the value of the kernel. Thus when classifying a new company, the values of its financial ratios are compared with the ones of the support vectors of the training sample, which is more similar to this new company. This company is then classified according to with which group it has the greatest similarity.

4. Experimental Results & Analysis

Our steganalytic algorithm proves 90% accuracy in correct detection of stego/normal image, when compared to an existing technique. The graph considers true and false positive attributes to show the correct detection as shown in Table 1. Table 1: is presented via graph in Fig 2.

Table 1: Correct detection accuracy is given and is compared with the existing system

Techniques	True Positive	False Positive	Correct Detection Rate
Ensemble Classifier	0.6	0.4	80%
SVM Classifier	0.8	0.2	90%

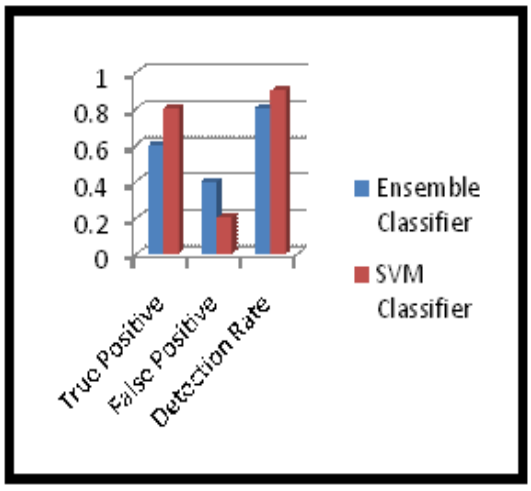

Fig 4: The proposed technique shows high correct detection rate

5. Conclusion

In this paper, we present a steganalytic algorithm based on residual value, which detects the stego/normal image with 90% accuracy. The accuracy rate remains stable even when different sets of images are tested. SVM classifier is employed over here, to classify between the images. In future, this potential work can be extended to video and the other directions can be considered.

References

[1] J. Fridrich, M. Goljan, and D. Hogea, "Steganalysis of JPEG images: Breaking the F5 algorithm," in Proc. Information Hiding, 5th Int. Workshop, 2002, vol. 2578, Lecture Notes in Computer Science, pp.310–323.

[2] Westfeld A.: High Capacity Despite Better Steganalysis(F5-A Steganographic Algorithm). In: Moskowitz,I.S. (eds): Information Hiding 4th International workshop. Lecture notes in computer science, vol.2137. Springer-Verlag, Berlin Heidelberg New York(2001) 289-302.

[3] Fangjun Huang, Jiwu Huang and Yun-Qing Shi. New Channel Selection Rule for JPEG Steganography (2012).

[4] J. Fridrich, M. Goljan, and D. Soukal, "Perturbed quantization steganography with wet paper codes," in Proc. ACM Workshop Multimedia and Security, Magdeburg, Germany, Sep. 20–21, 2004, pp.4–15.

[5] C. Fontaine and F. Galand, "How Reed-Solomon codes can improve steganographic schemes," EURASIP J. Inform. Security, vol. 2009, pp.1–10, 2009.

[6] D. Schönfeld and A. Winkler, "Reducing the complexity of syndrome coding for embedding," in Proc. Information Hiding, 9th Int. Workshop, 2007, vol. 4567, Lecture Notes in Computer Science, pp.145–158.

[7] Aura, T: Practical Invisibility in Digital Communication. In: Anderson,R.J. (eds): Information hiding: 1st International workshop. Lecture notes in computer science, vol.1174. Springer-Verlag, Berlin Heidelberg New York(1996) 265-278.

[8] J. Fridrich, "Feature-based steganalysis for JPEG images and its implications for future design of steganographic schemes," in Proc. Information Hiding, 6th Int. Workshop, 2004, vol. 3200, Lecture Notes in Computer Science, pp. 67–81.

[9] Westfeld , A. and Pfitzmann, A.: Attacks on Steganographic Systems. In:Pfitzmann A. (eds.): 3rd International Workshop. Lecture notes in computer science, vol.1768. Springer-Verlag, Berlin Heidelberg New York(2000) 61-75.

[10] Farid, H,: Detecting Steganographic message in digital images. Technical report, TR2001-412. Dartmouth College, New Hampshire (2001).
[11] C.-C. Chang and C.-J. Lin, LIBSVM:ALibrary for SupportVector Machines 2001 [Online]. Available: http://www.csie.ntu. edu.tw/~cjlin/ libsvm.
[12] D. Kundur and D. Hatzinakos, Digital Watermarking using Multiresolution Wavelet Decomposition, Proceedings, IEEE International Conference Acoustic, Speech, Signal Processing, 1998.
[13] Blossom Kaur, AmandeepKaur, Jasdeep Singh, Steganographic Approach for hiding Image in DCT Domain, International Journal of Advances in Engineering & Technology, July 2011.
[14] Nilanjan Dey, Anamitra Bardhan Roy, Sayantan Dey, A novel approach of color image hiding using RGB color planes and DWT, International Journal of Computer Applications, 2011.
[15] W. Zhu, Z. Xiong, and Y.-Q. Zhang, "Multiresolution watermarking for images and video: A unified approach," in Proc. ICIP98, Chicago, IL, Oct. 1998.
[16] VaishaliV.Jadhav, P.P.Belagali, Sapana Kishor Soudagar,Pooja Adgonda Patil, "Edge Adaptive Image Steganography using LSB Matching Revisited" IOSR Journal of Electronics and Communication Engineering (IOSR-JECE) ISSN : 2278-2834, ISBN : 2278-8735, PP : 40-44 .

T. J. Benedict Jose joined the faculty of the Department of Computer Science, St. Xavier's college as Assistant Professor. He is pursuing Ph.D. in Computer Science. He has published a paper in RAC Journal of research volume 2, which is a multi-disciplinary research journal.