AN OVERVIEW OF FEATURE BASED STEGANALYSIS

Deepa D. Shankar, Banasthali University, India
sudee99@gmail.com
Vinod Kumar Shukla, Amity University, UAE
vshukla@amityuniversity.ae

ABSTRACT

The field of cyber technology has grown tremendously over the years, which has made the transfer of data easy and fast. Open data can be of any form - text, image, audio, video and so forth. In order to transmit data through internet, it has to be converted into digital form. The advent of the internet has given people unlimited access to the digital world, which in turn may lead to huge security issues. Hence, the users need to protect their data from misuse. This has led to the birth of cyber security concepts like "data security in digital communication, copyright protection of digitized properties, and security of invisible communication over digital media". Most of the research on steganalysis has been done separately on targeted and blind steganography with various steganographic schemes. Different steganalysis techniques have been used, but statistical steganalysis scheme proved to be more accurate than any other. For targeted steganalysis, the feature selection and extraction are easier since the steganographic algorithm is known and the features that would clearly project the changes can be identified. For blind steganalysis, the distinguishing features need to be identified, selected and extracted based on the image format and the transformation of images.

Keywords: Steganography, Steganalysis, Cryptography, Message Format, Feature Analysis, Statistical analysis, Targeted Steganography, Blind Steganography

INTRODUCTION

Steganography is the science of obscuring data (audio and video) so that none, other than the sender and the probable recipient is aware of the hidden data. The object of the communication is the hidden message known as stego image, and the transmitting medium is called the cover image. Ssteganography is functional .The research on statistical steganalysis was initially done on transform domain by the means of blind steganalysis (Deepa D. Shankar, Gireeshkumar T., Hiran V.Nath,2011),(Kharrazi, M., Sencar, H.T., Memon, N.,2006), (Natarajan Meghanathan,Lopamudra Nayak,2010). The features such as mean, variance, kurtosis and skewness were calculated. These were the basic statistics used. The concept of calibration had come to light with the breaking of F5 algorithm using JPEG image formats (Fridrich, J., 2004). The calibration technique helped in detecting the embedding percentages as less as 10 %. Fridrich J.et al extended the study generally to other steganographic schemes (Pevn'y, T., Fridrich, J., 2007) & (Kodosky, J., Fridrich, J, 2009). The image format used was JPEG (Fridrich, J., Goljan, M., Hogea, D., 2003). They had used calibrated images to find the statistics. This was compared with F5, Outguess and Model based steganographic schemes. The statics used were of first order and second order. Fridrich J. et al later used a combination of DCT and Markov functions (Pevn'y, T., Fridrich, J., 2007). The calibrated techniques were later enhanced to work for steganographic algorithms like YASS (Jessica Fridrich, Miroslav Goljan, Dorin Hogea, David Soukal, 2003). PEV 274 feature set was used for this purpose. Pixel value differencing was used as a steganographic scheme by Tsai et al (Wu, D.C., Tsai, W.H., 2003). The block dependency scheme was tried on uncalibrated images by Deepa D. Shankar (Deepa D. Shankar, T.Gireeshkumar, K.Praveen, R.Jithin, Ashji S.Raj, 2012). This scheme was tried on different percentages of embedding and a combination of all. The differentiation and classification was done by means of support vector machines (Cristianini, N., Shawe-Taylor, J., 2000). The concept of principal component analysis is used to remove the redundant features (Miranda, A.A., Le Borgne, Y.A., Bontempi, 2008).

DIFFERENT TYPES OF ATTACKS IN STEGANALYSIS

Hiding data within electronic medium causes changes in the properties of the medium which can lead to degradation in some form. All images have some relation between neighboring pixels. There are different types of attacks in steganalysis. A few of them are described below.

Visual Attack: This method involves examining the images visually, for instance, scrutinizing bit images and trying to discover the difference visually. A change in the media may result in degradation of the quality of the media.

Structural Attack: The format of the cover image deviates when the data to be hidden is embedded. Understanding and identifying these typical changes can determine the existence of a hidden data.

Statistical Attack: In these attacks, hidden data is distinguished by the statistical analysis of the images by some mathematical principles. The hidden data is more haphazard than the original data. Therefore, finding a procedure to detect the randomness exposes the presence of data. Statistical investigations unfold the hidden data by defining that an image's statistical properties digress from the norm.

STEGANALYSIS ON IMAGES

Over years, pronounced development in fields of steganography and steganalysis has been observed. . Steganalysis is gaining importance and momentum in the field of computer forensics for selection and tracking of documents that are suspected of unregulated events for information security system to avert the leak of unapproved data through internet. Quite a few new steganography methods are being recommended each year, most of which are followed by improved steganalysis methods for their exposure. Steganalysis is a stimulating arena in the field of cyber security due to the dearth of information on the definite features of the cover image that is used to hide data and retrieve it later (Kharrazi, M., Sencar, H.T., Memon, N.,2006). Steganalysis can be divided into two - blind steganalysis and targeted steganalysis. This segment of steganalysis is intended for a known embedding algorithm. The distinguishing features needed for analysis is identified from the knowledge of the steganographic algorithm used Blind steganalysis are the scheme that is not dependent on a specific embedding method (Jessica Fridrich, Miroslav Golan, Dorin Hogea, David Soukal, 2003). It eliminates the dependency on the performance of individual embedding methods. Hence, this technique can work in a broad spectrum of steganographic techniques.

DCT DOMAIN STEGANALYSIS

It is presumed that the quantized DCT coefficients are tough for small distortions and after cropping the calculated DCT coefficients, it will not reveal clusters due to quantization (Fridrich, J.2004). Since the cropped stego image is visually like the cover image, many macroscopic features of the cover image will be roughly conserved. After forecasting DCT coefficient’s histogram in the original image and linking with that of a stego image, the hidden image length can be projected.

SPATIAL DOMAIN STEGANALYSIS

In statistical steganalysis, a steganalyst scrutinizes the embedding algorithm and then finds a few characteristic statistics those deviates with embedding, but recovers by means of calibration (Fridrich, J., Goljan, M., Hogea, D. 2003),(Fridrich, J.,2004), (Pevn’y, T., Fridrich, J.,2007). For JPEG images, calibration is done by decompressing the stego image, cropping a few pixels horizontally and vertically, and recompressing back using the same quantization table (Deepa D. Shankar, Gireeshkumar T., Hiran V.Nath, 2011). The distinctive statistics calculated are used as evaluation for the cover image.

FEATURE ANALYSIS

Blind steganalysis is composed of two principal components. These are feature extraction and feature classification (Kodosky, J., Fridrich, J, 2009). In feature extraction, a set of relevant statistics is identified and extracted. These features are suggested by detecting the image features that exhibit high variation when embedded with data. Feature classification, which is the next component, operates in two modes. The acquired statistics from both stego and cover images are first used to train a classifier (Cristianini, N., Shaw-Taylor, J., 2000). Before the features are fed into the classifier, features are reduced using principal component analysis (Miranda, A.A., Le Borgne, Y.A., Bontempi, 2008). Previous literature (Fridrich, J., Goljan, M., Hogea, D. 2003), state only the application of JPEG images in the either DCT domain or the spatial domain in terms of embedding and extraction. Feature-based steganalysis (Deepa D. Shankar, Gireeshkumar T., Hiran V.Nath, 2011), (Pevn’y, T., Fridrich, J., 2007) is a technique where certain features that are sensitive to embedding changes but insensitive to image content are extracted. It can be done in both calibrated and uncalibrated images.

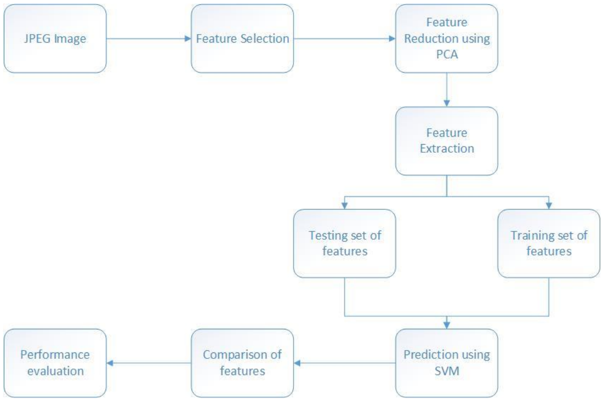

Figure 1: System Architecture for feature-based Classification scheme

CALIBRATION STEGANALYSIS

This steganalysis is usually used as a part of statistical steganalysis, in a transformed domain. This approach helps the blind steganalysis scheme since it has no idea of the steganographic algorithm. Calibration exploits the fact that the embedded payload in a stego image may be probably removed, when changes are made to the image, thus creating an equivalence of the cover image (Kodosky, J., Fridrich, J, 2009). The changes can be by means of cropping, shearing, and reducing the resolution.

The process of calibration is as follows:

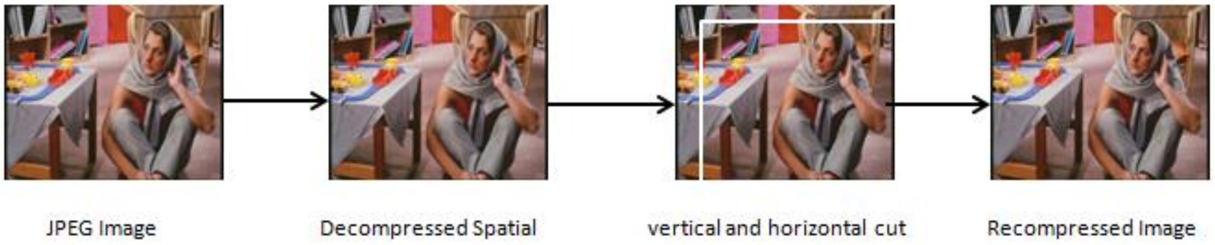

Figure 2 : Calibration Technique

During calibration, a transformed image (preferably DCT) is decompressed to its equivalent in spatial domain. The image is then cut horizontally and vertically by n pixels (the result is excellent when n=4). The image is then compressed back using the same parameters used before, i.e., quantization matrix, quality factor, etc.

QUANTITATIVE STEGANALYSIS

Quantitative steganalysis is used as a part of targeted steganalysis (Pevny, T., Fridrich, J., Ker, A.D., 2009). This approach is used to detect an estimate of the changes in embedding. This steganalysis scheme uses the embedding algorithm and identifies its effect on statistics, when a payload is embedded, and then exploits it. Since this is intended for targeted steganalysis, no training of a classifier is needed.

PRINCIPAL COMPONENT ANALYSIS

Principal component analysis (PCA) is used as a statistical procedure to identify the principal variables from a set of variables. In statistical steganalysis, the extracted features may have non-relevant ones as well, hence, principal component analysis is used to identify the relevant features and reduce the dimension with only the distinguishing features. The PCA had been used before to reduce the Markovian feature dimension from 324 features to 123 distinguishing features (Deepa D. Shankar, T.Gireeshkumar, K.Praveen, R.Jithin, Ashji S.Raj, 2012).

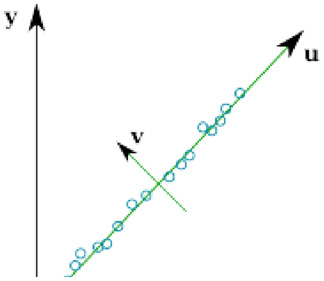

Figure 3: Principal Component Analysis for dimension reduction

SUPPORT VECTOR MACHINE

Support Vector Machine (SVM) in machine language is an algorithm, which uses supervisory technique for classification of data. Given a set of data for training, the SVM would output an optimal hyper plane which would clearly classify the data. Support vectors are datasets which lies closest to the hyperplane. These points are very difficult to classify. Hence they are able to change the position of the hyper plane. The hyper plane can be so decided to give the largest minimum distance to the support vectors. This distance is called margin.

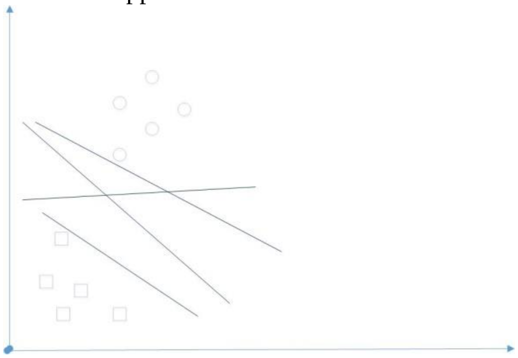

Figure 4: Classified data with different hyper plane

The classification can be done in many ways as shown in the figure above. But if the classification hyper plane is too close to a sample feature, it will be noisy and the classification will not be proper. Hence the hyper plane should be so selected that the line should be far from all the points and also should classify. Such a hyper plane is called optimal hyper plane (Tristan Fletcher Support Vector Machines Explained, UCL, UK).

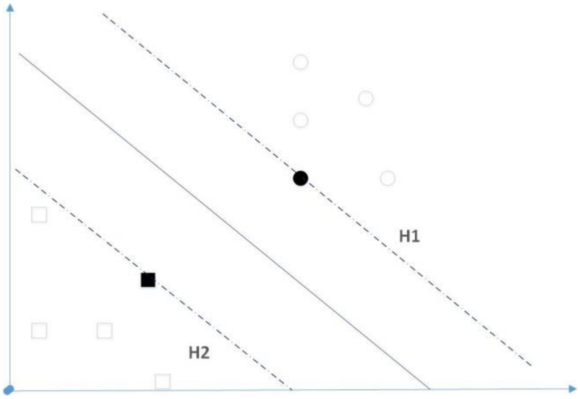

Figure 5: Classified data with optimal hyperplane in SVM

Support vector Machine works best in high dimensional spaces. The algorithm is versatile in the sense that many kernel functions can be tried in Support Vector Machine. The flexibility of this trial is less in other classifiers.SVM would also classify even if the number of dimensions is greater than the training dataset. However, the efficiency may be affected if the number of features is very much greater than the training dataset.

FUTURE WORK AND OTHER POSSIBILITIES

Very little research has been done in steganalysis so far, that incorporates machine learning techniques. Until now, and the research was done only in analyzing the image. This is useful in determining whether the image contains hidden data or not. Statistical steganalysis is found to produce better results when machine learning techniques were used for classification (Cristianini, N., Shawe-Taylor, J., 2000). The literature survey done so far has not revealed a comparative study of blind and targeted steganalysis (Fridrich, J.,2004), and has not produced a substantial result. Moreover, there are very few studies done on the comparison of the effect of calibrated and non-calibrated images in statistical steganalysis. The effect of the use of a transform technique in steganalysis and its comparative study too has not been done yet. The proposed research, represented in Figure 6 will fill the above mentioned gaps and concentrate on the steganalysis in different domains like spatial domain and transform domain, for both blind and targeted steganalysis. The idea is to compare the accuracy rate of statistical analysis for different steganographic schemes and different transformations, and hence identify the features that give an efficient analysis rate in all typical scenarios.

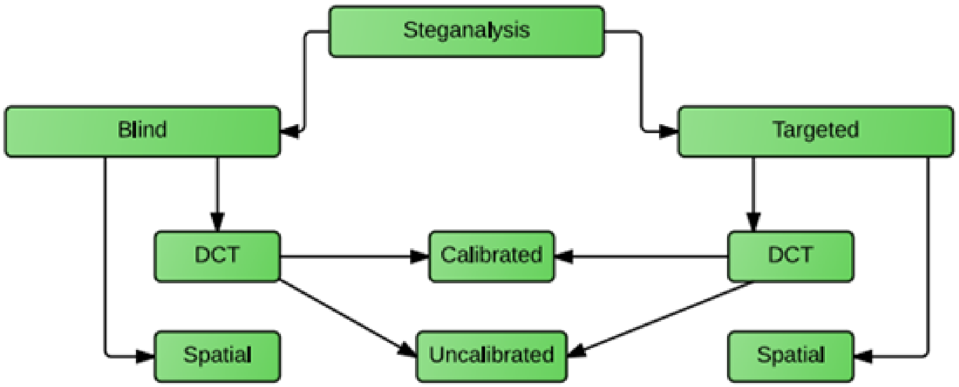

Figure 6 :Diagrammatical representation of the scheme of work

It has been assumed that lower message length will not yield result in DCT domain. Research has been done on embedding percentages till 25 (Pevn'y, T., Fridrich, J., 2007), (Tomas Pevny, Jessica Fridrich, 2005). Although the embedded message will be at least 25 % in real-time, to rule out any possibilities the message length of approximate 10 % in both DCT and spatial domain, has been taken up for research here.

CONCLUSION

Since steganalysis is a fairly new way to analyze the presence of hidden data, it has a long way to go. But till now statistical analysis has given better results than any other means. Once the results are evaluated, they will be compared with each other for the best outcome. The research aims at extraction of the best features, both in spatial and transform domain. The features would be extracted with calibrated images and uncalibrated images, since there is no assurance of the cover images in blind steganalysis.

REFERENCES

Cristianini, N., Shawe-Taylor, J. (2000) An Introduction to Support Vector Machines and other Kernel-Based Learning methods. Cambridge University Press.

Deepa D. Shankar, Gireeshkumar T., Hiran V.Nath (2011) Steganalysis for Calibrated and Lower Embedded Uncalibrated Images full paper LNCS Springer 7077 ISBN 978-3-642-27241-7 pp. 294-301.

Deepa D. Shankar, T.Gireeshkumar, K.Praveen, R.Jithin, Ashji S.Raj (2012) Block Dependency Feature Based Classification Scheme for Uncalibrated Image Steganalysis LNCS Springer 6411 pp-189-195.

Fridrich, J., Goljan, M., Hogea, D. (2003), Steganalysis of JPEG Images: Breaking the F5 Algorithm. In: Petitcolas, F.A.P. (ed.) IH 2002. LNCS, vol. 2578, pp. 310–323. Springer, Heidelberg.

Fridrich, J. (2004) Feature-Based Steganalysis for JPEG Images and Its Implications for Future Design of Steganographic Schemes. In: Fridrich, J. (ed.) IH 2004. LNCS, vol. 3200, pp.67–81. Springer, Heidelberg

Jessica Fridrich, Miroslav Goljan, Dorin Hogea, David Soukal (2003) "Quantitative Steganalysis of Digital images: estimating the secret message length" Multimedia Systems, Springer- Verlag 2003 DOI 10.1007/s00530-003-0100-9,Vol 9,Issue 3,pp 288-302 , 2003.

Kharrazi, M., Sencar, H.T., Memon, N. (2006) Performance study of common image Steganography and Steganalysis techniques. Journal of Electronic Imaging 15(4), 041104.

Kodosky, J., Fridrich, J (2009) Calibration Revisited. In: ACM Multimedia and Security
The workshop, Princeton, NJ, vol. 8, pp. 63–74.

Miranda, A.A., Le Borgne, Y.A., Bontempi (2008),New Routes from Minimal Approximation Error to Principal Components., Neural Processing Letters 27(3).

Natarajan Meghanathan,Lopamudra Nayak (2010),Steganalysis Algorithms for detecting the hidden information in Image, Audio and Video Cover Media,International Journal of Network Security and Application(IJNSA),Vol.2,No.1.

Pevn’y, T., Fridrich, J. (2007) Merging Markov and DCT Features for Multiclass JPEG
Steganalysis. In: Proceedings SPIE, Electronic Imaging, Security, Steganography and Watermarking of Multimedia Contents IX, San Jose, CA, vol. 6505, pp. 301–314.

Pevny, T., Fridrich, J., Ker, A.D. (2009) From Blind to Quantitative Steganalysis, SPIE, Electronic Imaging, Media Forensics and Security XI, San Jose, CA, January 18-22, vol. 14, pp. 0C1– 0C14.

Tristan Fletcher Support Vector Machines Explained, UCL, UK

Tomas Pevny, Jessica Fridrich, (2005) Towards Multi-Class Blind –Steganalyzer for JPEG images International Workshop on Digital Watermarking LNCS vol. 3710, Springer-Verlag, pp 39-53

Wu, D.C., Tsai, W.H. (2003) A Steganographic Method for images by pixel value differencing. Pattern Recognition Letters 24, 1613–1624.