Volume 3, Issue 12, December 2013 ISSN: 2277 128X

International Journal of Advanced Research in Computer Science and Software Engineering

Research Paper

Available online at: www.ijarcsse.com

Reliable LSB Steganography based on Variable Key Encryption

Mrs.K.Rajasri, Ms.T.Indhumathi
M.Tech, *Department of CSE,*
Christ College of Engineering and Technology, India

Abstract— This The only responsibility of steganography is to cover the piece of information so as to conceal any communication that is taking place. Steganalysis is the means used to notice the existence of any concealed communication in a cover medium. A fresh approach based on feature mining in the discrete cosine transform domain based machine learning for steganalysis of JPEG images is projected. The neighbouring joint density and features on both intra-block and inter-block are extracted from the DCT coefficient array respectively. Least Significant Bit based steganography is one of the simplest techniques that hide a clandestine communication in the LSBs of pixel values without introducing a lot of perceptible distortions. But, LSB method can easily be identified. So, we plan to combine the DCT features with the LSB embedding. If the secret data is in text format, then it is converted into mixed case font and then embedded in LSB. After the feature space has been constructed, it uses SVM like binary classifier for training and classification. After the training is finished the characters from test images are given for classification. The performance of this method on different Steganographic systems named F5 is analyzed. Individually each attribute and combined features categorization correctness is checked and concludes which provides better classification.

Keywords—Information Hiding, Steganography, Steganalysis

I. INTRODUCTION

The significant constituents of today's information hiding are cryptography, watermarking and steganography. Still, each of them has diverse objectives for serving this intention. Cryptography is the learning of processing digital data by scrambling or encrypting in data bits with a key in such a way that the information is unintelligent to the illicit individual who does not own the key to recover or decrypt it. Encryption encodes the information into an unreadable design called cipher so that an unintended receiver cannot decide its intended implication. [8] It is very clear in cryptography that the encrypted data stored in the memory or being transmitted takes unreasonable quantity of computer processing resources and time during its useful life time to decrypt it. Though, message data after decryption may at all time is dispersed in plain form without any constraint, even by the authorized client. In addition encryption clearly marks a message as containing interesting information, and the encrypted message becomes focus to attackers.

Watermarking of digital data, on the other hand is the procedure that enables data called a watermark, digital signature, tag, or ticket into a multimedia article such as text, audio, image or video in perceptually undetectable or impossible to hear way with no degrading of worth of the object, such that watermark can be detected or extracted soon after to make an affirmation about the object. The embedded information can be a sequential numeral or random number sequence, ownership identifiers, copyright messages, control signals, transaction dates, information about the creators of the work, bi-level or gray level images, text or other digital data formats. A significant objective of watermarking is to make removal of the inserted watermark bits from the watermarked object not possible with no degrading the feature of the object and with no supplementary information such as a key. Second important goal of watermarking is to sense that the object has been tempered by checking that the watermark is being removed or damaged. Third goal of watermarking is prevention against copying and transmitting music, image, video on CDs and DVDs. Violation of copyrighted materials such as music and video happens often. There has been no technique so far developed that meets the hope of watermarking as wanted. Also, it has been converted into a legal to develop, sell or dispense code-cracking commercial software and hardware devices for anti-piracy ways with the advent of Digital Millennium Copyright Act of 1998. Thus music and video industries no longer depend on watermarking to verify breach of DMCA for copyrighted materials, but they are now rely on other approaches such that, their Internet providers to locate the possible violators. Steganography is the art of concealed writing and clandestine communication. Steganalysis involves extraction of statistical measurements based on marginal and joint DCT statistics from clean and stego images.

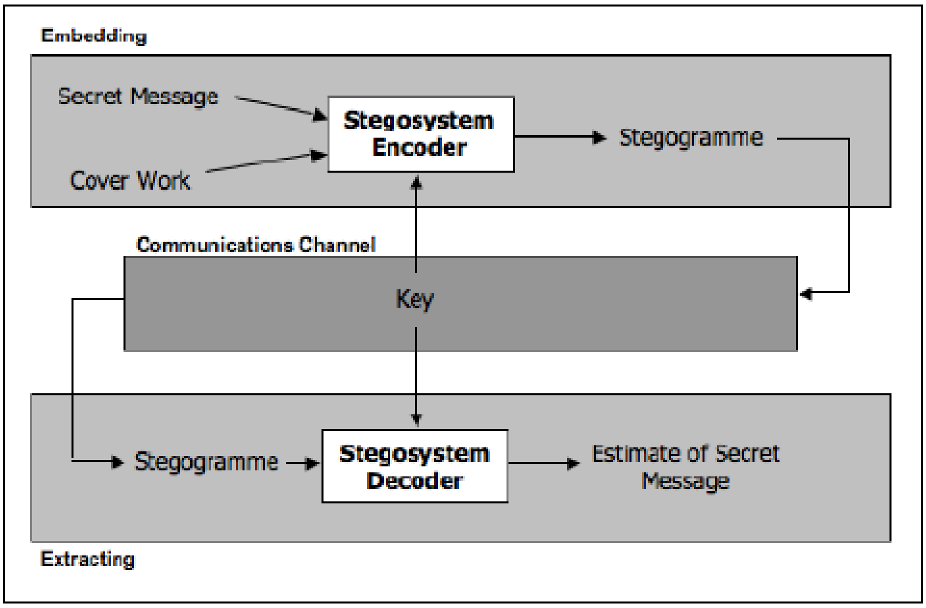

Fig. 1: Basic steganography system [5]

In this paper, we use LSB to hide message in the image.

The organization of this paper is as follows. In the next segment some previous DCT based steganalysis methods and LSB methods are summarized. Segment III is dedicated to the description of the existing method. Segment IV is devoted to the design of the proposed steganalytic scheme. In this section, we first initiate our motivation and then present mixed font method used to secretly embed the data. The conclusions are drawn in Segment V.

II. LITERATURE SURVEY

A. *Simple LSB Substitution*

Steganography is a procedure that involves hiding a communication in a suitable carrier for example an image or an audio file as in [1]. The most frequent and easiest method of concealing clandestine information is LSB substitution process. The main idea of LSB is to hide the clandestine in the least significant bit of the image. For example, for hiding image E into image H which we assume both of them to be n-bit gray scale images, a suitable method is to hide E into the position of the least significant bit of H. At first, the right most k LSBs from each pixel of H are extracted to create a k-bit gray-scale image, R, called the residual image which equals E in size. After that, E is altered to a k-bit image by disintegrating each pixel of E into various small k-bit units and processing each unit as a single pixel. The outcome of this disintegration is E'. Finally, R is replaced by E'. The embedding outcome is Z.

Different approaches have been employed to notice stego images. The first one consists in studying directly DCT coefficients like J. Fridich who looked at first order statistics and at the discontinuity of DCT coefficients at the limits of blocks for detecting the use of F5 and Outguess. She also pointed out some other characters for the frequency domain for JPEG steganalysis. The second approach is devoted to the spatial domain. The classifier with a high detection rate by combining Support Vector Machines with higher order statistics or with wavelet transform statistics of decompressed JPEG image is obtained. The novel means is projected to employ a Fisher discriminant as a replacement for SVM. Yet another method is introduced metrics based on images quality. Previous methods have even been used together to increase the correctness of detectors. Specific steganalysis is dedicated to only a known embedding algorithm. It may be very accurate for detecting images embedded with the given steganographic algorithm but it fails to detect those embedded with another algorithm.

Universal steganalysis enables to recognize stego images whatever the steganographic system be used. Universal steganalysis is also known as blind steganalysis. It is more appropriate and sensible than the specific steganalysis. This is for the cause that it can notice a larger class of stego images; it is usually less accurate for one given steganographic algorithm.

B. *Steganalysis of LSB Encoding*

The logic behind this method is that the LSBs in typical scanned images or images taken with a digital camera are fundamentally arbitrary, and replacing them with an encrypted message will not introduce any obvious artifacts as in [10]. LSB-based steganography, in which the lowest bit plane of an image is used to convey the secret information, has been used by many Steganographers, because the eye cannot detect the very small perturbations this means introduces into an image and also it is extremely simple to execute as in [3]. By means of extracting features from spatial and DCT domain, these steganalytic techniques had a good outcome for BMP images, including spatial domain and DCT domain hiding methods. Classifiers are generally used for distinguishing original and stego images. There are several steganalyzers used for this classification. Simpler classifiers are naturally less prone to over fitting. A three back propagation Neural Network is used as a classifier for identifying stego images as well as images. Non linear Support Vector Machine classifier is also used for effective classification of stego images and cover images. The steganographic algorithm YASS is an instance of a completely novel idea for constructing stego methods.

YASS is a JPEG steganographic method that hides data in the discrete cosine transforms coefficients of randomly chosen image blocks. It challenges the long recognized idea that an essential term for a safe technique is minimization of the embedding impact. Instead, YASS embeds the message strongly in a key-dependent domain and stores the embedded image as JPEG. The recompression at the same time with the key-based embedding domain, introduce large challenges for the steganalyst. The fact that the embedding changes are restricted in a domain that is protected away from the attacker by a key means that the steganalyst cannot easily recognize quantities that are distorted by embedding and those that stay invariant. Moreover, the impact of embedding changes is added mask by the recompression. Combined, these two design elements make steganalysis of YASS especially firm. In particular, steganalysis methods that use calibration are rather unproductive in detecting YASS because its embedding means does not in a straight line use the quantized DCT coefficients.

The scheme on which YASS is based is quite general and modular in the sense that the specific application may proceed in other domains, such as the wavelet domain, and the robust embedding algorithm may involve other embedding principles than QIM in a subband of block DCT coefficients. The novel technique is planned to steganalyze YASS using blind attacks that do not take advantage of any fault of a specific execution of the embedding algorithm. These attacks can be applicable to upcoming variants of YASS that would not suffer from lack of randomization.

Feature selection becomes very necessary for machine learning responsibilities when facing high dimensional data these days. Feature selection is used as a pre-processing step to machine learning. It is efficient in reducing dimensionality, removing unrelated data, increasing learning accuracy, and getting better consequence clarity. However, the recent increase of dimensionality of data poses a strict challenge to a lot of existing feature selection methods with respect to efficiency and efficacy. The novel concept of predominant correspondence is proposed. A fast filter is used. It can distinguish relevant characters from redundancy among relevant descriptions without pair wise correlation analysis. The efficiency and efficacy of this means is demonstrated through extensive comparisons with other methods using real-world data of high dimensionality. On the other hand, the style of extent on both size and dimensionality also poses severe challenges to feature selection algorithms. Various new research efforts in feature selection have been focused on these challenges from using a huge number of instances for dealing with high dimensional data. The work about feature selection for high dimensional data is concerned.

Feature selection algorithms fall into two broad categories, the filter model or the wrapper model. The filter model relies on general features of the training data to select some features without involving any learning algorithm. The wrapper model requires one predetermined learning algorithm in feature selection and uses its performance to assess and decide which features are selected. As for each new subset of features, the wrapper model needs to learn a hypothesis. It tends to find out descriptions better suited to the predetermined learning algorithm resulting in superior learning performance. But it also tends to be more computationally expensive than the filter model. When the number of features becomes very large, the filter model is frequently selected due to its computational efficiency.

III. EXISTING SCHEME

A. *Ensemble Classification for Steganalysis*

The ensemble:

In order to boost the mutual range of the base learners, each learner is trained on a bootstrap model drawn from the training set rather than on the entire training set as in [6].

B. *Optimal Pixel Adjustment*

The hiding procedures are shown in the next steps as in [2]. There are many versions of spatial steganography. All those methods directly alter some bits in the image pixel values in hiding data. Least significance bit-based steganography is one of the simplest techniques that conceal a clandestine message in the LSBs of pixel values without introducing many visible distortions. To our human eye, changes in the value of the LSB are invisible. Thus, LSB can be used as a perfect place for hiding information. This does not engage any perceptual alter in the cover object. Embedding of message bits can be done either in sequence or at arbitrary.

C. *Relation to other Existing Steganalysis Methods*

Though sample-pair analysis is not restricted to adjacent pixels, the authors note the estimate is better in training for spatially neighbouring samples as in [4]. The image is disintegrated to a set of binary images according to the bit-plane complexity segmentation, which divides bit-plane into consecutive and non-overlapping blocks. Each block is further checked whether it is noise-like or not, and noise-like blocks are appropriate for embedding information.

JPEG is based on DCT in lossy compression and it is the most ordinary arrangement of images produced by digital cameras, scanners and other photographic capture devices. In JPEG compression, consecutive sub image blocks of size of 8X8 on applying DCT produces 64 DCT coefficients. Data can be inserted in these coefficient’s insignificant bits. However, changing any single coefficient would influence the whole 64 block pixels. No visible change can be seen in the stego image as the changes due to insertion of data are in frequency domain. JSteg embeds secret message into a cover image by consecutively replacing the LSBs of non-zero DCT coefficients with message bits. Then PSNR value can be calculated using PSNR = 10 log 10MAX2/MSE (db).

LSB Steganography can be classified by two methods LSB substitute and LSB matching. The LSB embedding causes the occurrence of individual requisites of a Pairs of Values to flatten out with respect to one another. So for example if an image has 100 pixels that have a value 2 and 200 pixels that have a value 3, then after LSB embedding of the entire LSB plane the accepted frequencies of 2 and 3 are 150 and 150 respectively. This of course is when the whole LSB plane is modified. Though, as long as the embedded message is greatly sufficient, there will be a statistically visible flattening of Pairs of Value distributions and this fact is demoralized by their steganalysis method. An encryption algorithm has been anticipated earlier to steganography for invention of encoded message as in [9]. The main crisis of this technique is that, hiding the information in the channels is done in a logical method.

The version of this template is V2. Most of the formatting instructions in this document have been compiled by Causal Productions from the IEEE LaTeX style files. Causal Productions offers both A4 templates and US Letter templates for LaTeX and Microsoft Word. The LaTeX templates depend on the official IEEEtran.cls and IEEEtran.bst files, whereas the Microsoft Word templates are self-contained. Causal Productions has used its best efforts to ensure that the templates have the same appearance.

IV. PROPOSED SCHEME

Whenever the secret message to embed is in text form, then the message is transformed in the mixed-case font and then encrypted before embedding in the cover medium.

A. *JPEG domain rich model*

The simplest and most popular method in special domain is least significant bit which embeds data in the LSB position of bit pixel of a cover image. A simple method for embedding a message into a digital image is to alter the least significant bit of the image pixels, so that the LSBs of consecutive pixels encode a message. In so doing, the perceptual distortion to the cover image is nearly insignificant and unlikely to be detected by simple visual inspection. It is not surprising that with the emergence of steganography that the development of a counter-technology, steganalysis, has also emerged.

The mixed case font method is proposed along with LSB steganography. In this method, the text message is written in the format that the letters of the message are in upper and lower cases alternatively. For example, ‘Steganos’ can be written as ‘sTeGaNoS’. By this method, a large volume of information can be hidden in text compared to other methods, since the proposed technique does not use spaces among words or between paragraphs but use only the letters themselves. The goal of steganalysis is to decide if an image contains an embedded message. As this field has developed, determining the length of the message and the actual contents of the message are also becoming active areas of research.

B. *Building the rich model*

First, we model individual DCT modes individually. Then, gather a lot of these sub models and put them together. Current steganalysis methods fall broadly into one of two categories: embedding specific or universal. While universal steganalysis attempts to notice the existence of an embedded message independent of the embedding algorithm and, ideally, the image format, embedding specific approaches to steganalysis take benefit of particular algorithmic details of the embedding algorithm. Given the ever growing numeral of steganography tools, universal approaches are clearly essential in order to perform any type of generic, large-scale steganalysis. Specifically, we showed how a statistical model based on first- and higher-order magnitude statistics extracted from a wavelet decomposition, coupled with a linear discriminant analysis, could be used to detect steganography in gray scale images. Then it replaced the LDA classifier with a nonlinear support vector machine, affording better classification correctness. Then extended the statistical model to colour images, and described a one-class SVM that simplify the training of the classifier. In this Culminating paper, we summarize the results and expand the statistical form to comprise phase info.

Finally, the vital aim is achieved i.e. only sender and receiver know about the information.

V. CONCLUSION

This is a proficient approach to plot clandestine message into gray scale images to provide better image quality and information embedding capacity. Fixing the communication length allows us to compare the safety of dissimilar embedding schemes that might have a different capacity.

REFERENCES

[1] Manish Mahajan, Akashdeep Sharma, *Steganography in Colored Images Using Information Reflector with 2k Correction*, International Journal of Computer Applications (0975 – 8887) Volume 1 – No. 1, 2010.

[2] Parisa Gerami, Subariah Ibrahim, Morteza Bashardoost, *Least Significant Bit Image Steganography using Particle Swarm Optimization and Optical Pixel Adjustment*, International Journal of Computer Applications (0975 – 8887) Volume 55– No.2, October 2012.

[3] Sedighe Ghanbari, Manije Keshtegary, Najme ghanbari, *New Steganalysis Method using Glcm and Neural Network*, International Journal of Computer Applications (0975 – 8887) Volume 42– No.7, March 2012.

[4] Kenneth Sullivan, Upamanyu Madhow, Shivkumar Chandrasekaran, and B. S. Manjunath, *Steganalysis for Markov Cover Data With Applications to Images*, IEEE Transactions On Information Forensics And Security, Vol. 1, No. 2, June 2006.

[5] Abdelmgeid Amin Ali, Al - Hussien Seddik Saad , *New Text Steganography Technique by using Mixed-Case Font*, International Journal of Computer Applications (0975 – 8887) Volume 62– No.3, January 2013.

[6] Jan Kodovský, Jessica Fridrich,and Vojtech Holub, *Ensemble Classifiers for Steganalysis of Digital Media*.

[7] Siwei Lyu and Hany Farid, *Steganalysis Using Higher-Order Image Statistics*, IEEE Transactions On Information Forensics And Security, Vol. 1, No. 1, March 2006".

[8] Deepesh Rawat, Vijaya Bhandari, *Steganography Technique for Hiding Text Information in Color Image using Improved LSB Method*, International Journal of Computer Applications (0975 – 8887) Volume 67– No.1, April 2013.

[9] Souvik Bhattacharyya, Indradip Banerjee and Gautam Sanyal, *A Novel Approach of Secure Text Based Steganography Model using Word Mapping Method (WMM)*, International Journal of Computer and Information Engineering 4:2 2010.

[10] Bhanu Prakash Battula, R. Satya Prasad, *Essentials of Image Steganalysis Measures*, Journal of Theoretical and Applied Information Technology.