ISSN:0975-9646

Statistical Steganalysis Scheme of Using Block DCT

Papiya Chakraborty
Assistant Professor,
Computer Science and Engineering Department,
Pailan College of Management And Technology

Dr. Bikramjit Sarkar
Assistant Professor,
Computer Science and Engineering Department,
Dr. B.C. Roy Engineering College, Durgapur

Abstract- Steganalysis is the process of cracking to identifying steganography by scrutinizing different Statistical parameter of a stego media. First of all a mathematical analysis may disclose statistical discrepancy in the stego medium. After that steganalysis process determines whether that media contains hidden message or not and then try to recover the message from it. The information hiding process changes the statistical properties of the cover, which is a steganalyst attempts to detect is called statistical steganalysis. The enormous growth in public domain channels and communication technology (i.e. Internet) has greatly facilitated transfer of huge data. However, such open network channels have better susceptibility to security threats causing illegal information access. In this case both steganography and steganalysis have received a great deal of awareness from law enforcement and the media. Some promising methods for universal steganalysis with respect to features and arrangement procedure have also been recognized.

In my research work I want to develop an image based steganalysis that combines Discrete Cosine Transform (DCT), and compression techniques with LSB techniques on any raw image. Initially the cover-image is transformed from spatial domain to the frequency domain using DCT. The image is then quantized. and LSB technique is used to insert in pixels specified according to a range.The goal of this paper is to construct a new still JPEG steganalyzer with the help of DCT quantizer with markedly improved statistical performance.

Index Terms: Steganography, Steganalysis, Stego-system, DCT, LSB, quantizer, frequency domain

1. NEED FOR THE STEGANALYSIS

Scientific development that occurs should advance in an professional manner to uplift the society. Researchers wants to find the surplus materials which defects the growth of Technology. The development of efficient technology gets concealed by the progression of security threats. In order to hide information passed by military and government, steganography is one of the techniques widely used. Steganalysis passes between two phases a) Detection and b) Extraction and that too without significant algorithm used for thrashing the information.

Hence, steganalysis is an art and also a science whereas detection is an art of finding whether hidden message exists or not and extraction is the science of applying the powerful method to unhide the message. Apart from all modern sciences and technologies, Statistical Steganalysis plays a vital role in capturing and representing both linear and non linear relationships. This is an intelligent system which helps to enable machines solves problems like human by extracting and storing the knowledge. To incorporate intelligent method for steganalysis, statistical steganalysis used to overcome the drawbacks of the conventional methods. And it helps to restore the ststistical properties of hidden image. This research work, active ststistical steganalysis is focused trying to concentrate on detection of the presence of hidden content within the image with improved block co-efficient detection accuracy and performance using Discrete cosine transformation.

2. MOTIVATION

Recently in digital age there is more probability for altering the information represented by an image without leaving any traces of interfereing. In most of fields such as surveillance systems, criminal investigation, medical imaging, journalism and intelligence services, forensics investigation, need consistency while transferring the information in the form of an image. Planning is the crucial part and the information planning is passed to others through covert communication in order to hide from government and other people. The effective medium of hidden communication is achieved by steganography. We know that terrorists used steganography for secret communication since 11th September 2001 attack. Politicians use steganography communication to express their political thoughts which are more sensitive in this world.

Diversity of data embedding algorithms and different types of images makes the steganography a tough operation for researchers to develop a powerful technique for steganalysis.

3. INTRODUCTION ABOUT STEGANALYSIS

Steganalysis is a comparatively new research area with a small number of articles appearing before the late-1990s. Government officials, especially in the US, have shown a large interest on steganography and steganalysis research over the last years.

Many articles, were published at that time and brought the world's attention on the use of steganography. Most of the generating process of hiding data that were developed in the last years have been successfully attacked. The statistical properties of the hidden data that secure stego-systems should preserve and each time a more secure embedding algorithm is developed, steganalysts find a new statistic that they can identify their attack space. Most steganalysis publications are written in the last 10 years. Steganalysis is "the process of detecting steganography by looking at variances between bit patterns and unusually large file sizes" **[1].** Steganalytic algorithm is the process of breaking steganography algorithm When we will apply statistics on algorithm or by applying mathematical formula we can break the encryption algorithm but we can restore

the statistical part of the data that is called statistical steganalysis. Steganalysis Methods are broadly classified into three sub categories:

1.Supervised learning based steganalysis
2.Blind Identification based steganalysis
3.Parametric statistical steganalysis

In supervised learning based steganalysis, a classifier is constructed to differentiate stego and non stego images. Training inputs (both stego and non stego images) will be given to a learning machine. The classification rule is updated by a learning classifier based on prediction and ground truth. Finally, the stego classifier is obtained. This research focuses on supervised learning.

In blind identification based steganalysis, the statistical information is used to analyze images. Training data is not available. The specialty of this method is not only used to detect the presence of hidden information but also to extract it.

In Parametric statistical steganalysis, the detection is based on the available statistics. The statistical information may be completely known, partially known, or completely unknown. Hybrid technique combines more than one of the above mentioned methods

We can identify that inserted encrypted data into an image can changes the histogram of its color frequencies. As a result, the difference in color frequency between two and three has been a bit of removed by the inserted data. Instead of measuring the color frequencies, we analyze the frequency of the DCT coefficients. Steganalysis is a new approach to detect hidden messages from images made from stego system. In some cases we may know that the used steganographic system changes some properties of the medium with specific patterns on some of its properties. The techniques of this group try to detect the existence of steganographic data by detecting these patterns. Supervised learning based steganalysis. These techniques use statistical classifiers to see whether the tested image is a stego image.

As described Fig-1, in the following block diagram we can easily understand that using a secret key we can hide a secret information with the help of a cover image. All done by a process named Embedded process and an object was generated named stego object.(object similar as cover image with hidden information).To extract the message one should have this key. We can get the original secret message by applying extraction algorithm. The process of developing the extraction algorithm is called steganalysis

Basic Embedding and Extraction Process :

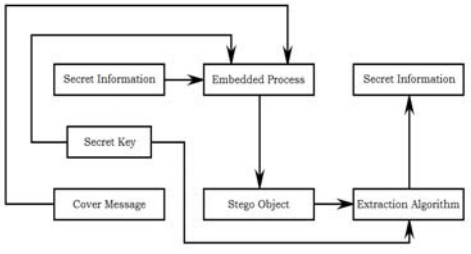

Fig-1

4. LITERATURE SURVEY

One of the most new and rapid growing technology is steganalysis. There has been an interest in understanding the distributions of the DCT coefficients since more than 20 years ago. after we have performed the DCT on each of the blocks and collected the corresponding coefficients from them, what will be the resulting statistical distribution? Such knowledge would be useful, for instance, in quantizer design and noise improvement for image enhancement [**2], [3**]. In this paper we have seen a typical plot of the histograms of the DCT coefficients.

Many Researchers are trying to locate statistical properties of images that the stego-system doesn't protect or find methods that one can find out if the image was altered at all or not. Neils Provos[**4, 5]** in 2001 used his steganalytic software StegDetect in order to test a large sample of images , downloaded from Usenet and eBay using a web crawler. He used a distributed dictionary attack on suspected stego images, but a very small percentages of the images were tested, and he wasn't been able to find any secret messages.

Neil F. Johnson and S. Jajodia [**7,8**] go for sorting steganalysis attacks to improve the message property, based on information on hand. With cryptography, assessment was made between any promising parts of the plaintext and parts of the ciphertext.. The message in a stego object may or may not be encrypted. If it is encrypted and the message is extracted, the cryptanalysis techniques may be applied.

In 2001 the RS (Regular/Singular) scheme was eshtablished by Fridrich, Goljan and Du [**6**] . Their approaches are counts the number of occurrences of pairs in sets. The idea is that spatially highly correlated adjacent image pixels, which can give us a lot of information whether LSB has been applied in the examined image, meaning that if LSB has been applied then areas where embedding has been made then adjacent image pixels would appear to have many different properties compared to where no tampering has been made.

Thus, steganalysis is considered successful if it can estimate whether an image contains a hidden message or not and the probability will be higher than random guessing. Steganalysis also attempts to find more information of the image and hidden message such as the type of embedding algorithm, the length of the message, the content of the message or the secret key used. A steganalysis attack can find any of the above and one can lead to another.

Jhonshon also describe Various types of techniques used for define attack in steganalysis . corresponding techniques are named as cryptanalysis. Attacks categories for cryptanalysis are *ciphertextonly*, *known plaintext*, *chosen plaintext*, and *chosen ciphertext*. [**7,8**] .

With the use of image redundancy Steganography method of digital image **[9, 10**] helps to hide the secret data. The methods are divided into two categories.They are spatial-domain methods and frequency-domain methos. In the spatial domain, the secret messages are embedded in the image pixels directly by LSB substitution method and Distortion technique, in which pixel property is changed

according to cover message and deflection of distorted image from original image contains secret information where In the frequency-domain, however, the secret image is first transformed to frequency-domain,with the help of FFT(Fast fourier Transform) and DCT(Discrete cosine transform) and then the messages are embedded in the transformed co-efficient accordingly **[11,12].**

5. OBJECTIVE AND THE SCOPE OF THIS RESEARCH WORK

The objective of this research work is to propose improved intelligent steganalysis:

1. To propose new block DCT algorithm for better identification of exposed information.
2. Mathematical and computational effort in learning cover and information images.
3. To make sure that an accurate detection of embedded information is possible when a forensic expert is interested in
4. These papers revise some innovative ways to enhance the steganalysis process in digital images.
5. The objective of this work is to develop and validate a fresh analysis method to provide Performance improvement over the various steganography methods in image frequency domain using discrete cosine transform co-efficient analysis using dual statistics method.
6. A histogram analysis using cosine parameter analysis the digital image enhance the encrypted secret bits in the special frequency domain which provides a model that meets both robustness as well as undetectable.

6 . RELATED WORKDONE:

a lot of experiment has been done on steganalysis. In the very first stage of the improving steganalysis work LSB steganography is formulated. There are Many steganalytic methods such as Chi-square statistical attack [14], RS analysis, sample pair analysis (SPA) analysis, weighted stego (WS) analysis, and structural steganalysis etc. But LSB steganography is one of the most successful method.

Another Type of steganalytic techniques [15] have been proposed in this paper. for example, the Chi-square attack, are effective to LSB steganography for spatial images as well as JPEG images. The fact that LSB steganography is susceptible to attack implies that high insensitivity does not give assurance a high security level. Still the first statistical steganalysis was proposed by Westfeld and Pfitzmann [16]. There approach is to define LSB embedding and is based on potential first order statistical analysis. It identifies Pairs of Values (POVs) that consist of pixel values, quantized DCT coefficients or palette index which get plotted to one another on LSB flipping.

After the message embedding, the total number of occurrence of two members of certain POV remains the same. This concept of pair wise dependencies is leads to design a statistical Chi-square test to detect the hidden messages [17].

Zhang and Ping [18] proposed a technique with the help of gray scale images. This technique employs different images histogram as the statistical analysis device. Measure of the weak correlation between the LSB plane and the rest of the planes is done by the translation coefficients between different image histograms. This algorithm can recognize the existence of secret messages embedded using sequential or random LSB replacement in images and also can calculate approximately the amount of secret messages.

Benton and Chu [19] proposed a soft computing approach to steganalysis specific to LSB For A better performance and computation speed than RS analysis method. In this paper Decision trees and neural networks are used independently for detection purpose. The features are extracted from images which are based on the variables for estimating the embedding probability in the RS method. This approach different from original RS method. The objective of this method is to choose whether the image contains hidden data but not to calculate the embedding probability. Xiang-dong Chen, et al. [20] a proposed a steganalysis technique based on bit plane randomness tests. Two binary sequences are obtained by scanning the 7the and 8th bit planes of the image with Hilbert scan. The randomness of these two sequences is tested individually by 14 kinds of randomness tests. The results of these tests form a vector and are used to construct a SVM classifier to distinguish stego images from the clean ones.

LSB matching more difficult and hard to detect as compared to simple LSB replacement. This study presents a survey of LSB matching steganalysis for digital image. Andew D. Ker et al. proposed a steganalysis technique for LSB matching in [21].

a steganalysis approach for both grayscale and color images described in H.B.Kekre et al. [22]. element vectors are calculated from gray level co-occurrence matrix (GLCM) in spatial domain has been used, which is sensitive to data embedding process. Different features of stego and non-stego images calculated and their differences is used for steganalysis. Absolute distance and Euclidean distance measurd for classification.

7. STEGANALYSIS METHOD FOR JPEG IMAGE APPLYING DCT

By changing the coefficients of transformation of an image data is implanted into the cover image, such as discrete cosine transform coefficients. It may be subjected to the lossy if we embed information in spatial domain in any image processing technique like compression, cropping etc. To overcome this problem we embed the information to be hidden in frequency domain. As the digital data is not continuous, to analyze the data of the image, we apply transformations to the image. We embed the data to be hidden by changing the values of the transformation coefficients accordingly. There are mainly three transformation techniques:

1. Fast Fourier transformation technique(FFT)
2. Discrete cosine transformation technique (DCT).
3. Discrete Wavelet transformation technique (DWT).

The main successful techniques are same in all three but our main attention in this paper is on JPEG images and they use DCT for steganalysis calculation. The information is hidden in the LSB's of the DCT coefficients of a JPEG image.

Every steganographic methods attempt to get the least amount of distortion in order to minimize the likelihood of introducing visible artifacts. In this experiment the cover-image, was initially stored in the JPEG format the spatial domain of embedding image will disturb but *not* erase the feature composition created by image encoding method.

Histogram of encoded image can still easily determine whether a given image has been stored as JPEG form. Indeed, it is possible to recover the JPEG quantization table from the stego-image by carefully analyzing the values of DCT coefficients in all 8×8 blocks.

After message embedding, however, the cover-image will become incompatible with the JPEG format in the previous image that it may be possible to prove that a particular 8×8 block of pixels could not have been produced by JPEG decompression of any block of quantized coefficients

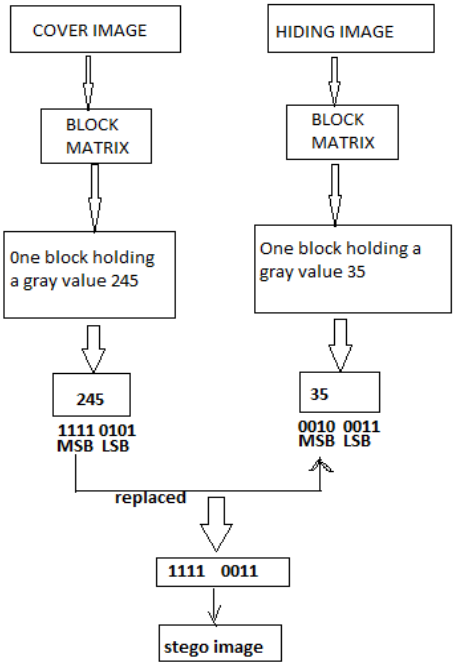

Fig-2

In Fig 2 we can see that 1st 4rth MSB of cover image is replaced with 1st 4rth MSB of hiding image and a new binary code 11110011 which is 243 in decimal is created, that is 99% same as cover image and this is called stego image. From that result we can confirm that the block is slightly modified. Indeed, it is highly doubtful to find an image stored in a lossless format . yet it is not fully compatible with any JPEG compressed image. By checking the JPEG compatibility of every block, we can potentially detect messages as short as one bit. And the steganalytic method will work for virtually any spatial steganographic method, not just the LSB embedding. One can even attempt to estimate the message length and its position in the image by determining which 8×8 blocks are incompatible with JPEG compression. It is even possible to analyze the image and estimate the statistical property for the cover-image or its blocks (the "closest" JPEG compatible image/block). This way, we may be able to identify individual pixels that have been modified.

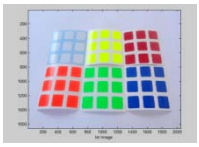

Fig-2 cover image

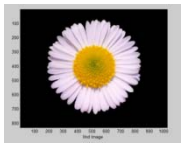

Fig-3 hiding image

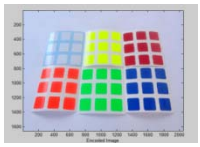

Fig-4 encoded image

Another difficult way of covering a secret data inside an image comes with the application and modifications of discrete cosine transformations. Discrete cosine transformations (DCT)), are used by the JPEG compression algorithm to transform successive 8 x 8 pixel blocks of the image, into 64 DCT coefficients each. Each DCT coefficient F(u, v) of an 8 x 8 block of image pixels f(x, y) is given by:

$$F(u,v) = \frac{1}{4}C(u)C(v)\left[\sum_{x=0}^{7}\sum_{y=0}^{7} f(x,y) * cos\frac{(2x+1)u\pi}{16} cos\frac{(2y+1)v\pi}{16}\right]$$

where C(x) = 1/√2 when x equals 0 and C(x) = 1 otherwise. After calculating the coefficients, the following quantizing operation is performed:

$$F^{Q}(u,v) = \left\lfloor \frac{F(u,v)}{Q(u,v)} \right\rfloor$$

where Q(u, v) is a 64-element quantization table[13]. A simple block diagram to find the difference between DCT co-efficient of an stego image i.e encoded image and an original cover image is look like this :

Block diagram to find whether an image contain any secret data or not

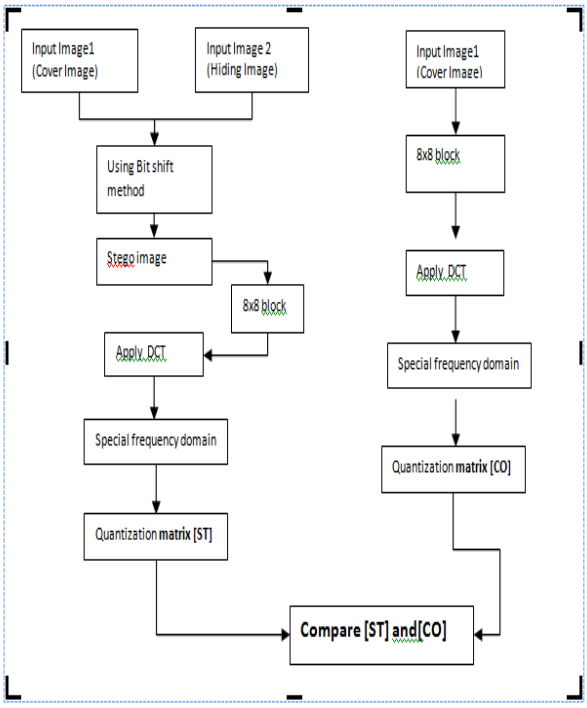

Fig-5

8. CONCLUSION

This sysnopsis presents the aim of this research work, Introduction of Steganalysis, need for the research work,Literature survey and motivation for the research work. Objective and the scope of the research work are also be presented. The extra data occurance within an image causes inconsistency in the image statistics, and the Steganalysis process determines whether the media contain any hidden message or not and try to recover it with preserving the image quality. Previous work utilizing first order statistics for steganalysis was fruitful, and introduced a lots of of algorithms that embedded data within the image while preserving the first order statistics.

Recent work will bring to a close that natural images also have meaningful second order statistics. This has been taken into consideration in recent steganalysis methods.

REFERENCE :

[1] Arvind Kumar, Km. Pooja, "Steganography-A Data Hiding Technique" International Journal of Computer Applications ISSN 0975– 8887, Volume 9– No.7, November 2010.

[2] G. Lakhani, "Adjustments for JPEG de-quantization coefficients," in *1998 Proceedings Data Compression Conference*, J. A. Storer and M. Cohn, Eds. Los Alamitos, CA: IEEE Comput. Soc. Press, Mar. 1998,p. 557.

[3] R. Brown and A. Boden, "*A posteriori* restoration of block transform compressed data," in *1995 Proceedings Data Compression Conference*, J. A. Storer and M. Cohn, Eds. Los Alamitos, CA: IEEE Comput. Soc. Press, Mar. 1995, p. 426.

[4] Provos, N. and Honeyman, P. (2002) 'Detecting Steganographic Content on the Internet',ISOC NDSS'02, San Diego, CA.

[5] Provos, N. and Honeyman, P. (2003) 'Hide and Seek: An Introduction to Steganography',Proc. IEEE.

[6] J. Fridrich, M. Goljan, and R. Du, Steganalysis based on JPEG compatibility, *Proc. of SPIE*, vol. 4518, pp. 275-280, 2001

[7] Johnson, N.F. and Jajodia, S. (1998) 'Exploring Steganography: Seeing the Unseen', George Mason University, IEEE Computers.

[8] Johnson, N.F. and Jajodia, S. (1998) 'Steganalysis of Images Created Using Current Steganography Software', Workshop On Information Hiding Proceedings, Portland, Oregon, USA.

[9] Kurak, C., and J. McHughes, "ACautionary Note On Image Downgrading," in IEEEComputer Security Applications Conference 1992, Proceedings, IEEE Press, 1992, pp.153-159

[10] Johnson, N. F., and S. Jajodia, "Exploring Steganography Seeing the Unseen, IEEE Computer , vol. 31, no. 2, 1998, pp. 26-34.

[11] Bender, W., D. Gruhl, and N. Morimoto, "Techniques for data hiding", IBM Systems Journal, vol. 35, no. 34, 1996, pp. 131-336.

[12] M¨oller, S., A. Pitzmann, and I. Stirand, "Computer Based Steganography How It Works and Why Therefore Any Restrictions on Cryptography Are Nonsense, At Best, in Information Hiding" First International Workshop, Proceedings, vol. 1174 of Lecture Notes in Computer Science, Springer, 1996, pp. 7-21.

[13] Papiya.Chakraborty "A Survey Analysis for lossy image compression using Discrete cosine Transform" International Journal of Scientific & Engineering Research Volume 3, Issue 9, September-2012 ISSN 2229-5518.

[14] Souvik Bhattacharyya,Indradip Banerjee and Gautam Sanyal, " A Survey of Steganography and Steganalysis Technique in Image, Text, Audio and Video asCover Carrier," Journal of Global Research in Computer Science, Volume 2, No. 4, April 2011,pp.1-15.

[15] Bin Li, Junhui He, Jiwu Huang, Yun Qing Shi, "A Survey on Image Steganography and Steganalysis," Journal of Information Hiding and Multimedia Signal Processing ,Volume 2, Number 2, April 2011 ,pp.142-172

[16] Westfeld, A.Pfitzmann, " Attacks on steganographic systems," Proc. of Information Hiding, Third Int. Workshop, Dresden, Germany, September 28–October 1, 1999, pp. 61–75.

[17] N.F. Johnson, S. Jajodia, "Steganalys of images created using current steganography software, in: Lecture Notes in Computer Science," vol. 1525, Springer-Verlag, Berlin, 1998, pp. 273–289.

[18] T. Zhang, X. Ping, "Reliable detection of LSB steganography based on difference image histogram," in: Proc. ICASSP, vol. I, 2003, pp. 545–548.

[19] Ryan Benton, Henry Chu, "Soft computing approach to steganalysis of LSB embedding in digital images," in: 3rd Int. Conf. on Information Technology Research and Education, 27–30 June 2005, pp. 105–109.

[20] Xiang-dong Chen, "Detect LSB steganography with bit plane randomness tests," in: Proc. of 6th World Congress on Intelligent Control and Automation, China, June 21–23, 2006.

[21] A.D. Ker, "Steganalysis of LSB matching in grayscale images," IEEE Signal Process. Lett. 12 (6), June 2005,pp. 441–444.

[22] H.B. Kekre, A.A. Athawale & S.A.Patki, " Steganalysis of LSB Embedded Images Using Gray Level Co- Occurrence Matrix," International Journal of Image Processing (IJIP), Volume 5 ,Issue 1: 2011.