ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 2, February 2014

Image Steganalysis Based on Statistical Evidance by Using SVM

Ms. Tejashree Shinde[1], Ms. Ujwala Chaudhari[2], Ms. Rushali Bodke[3]
UG Student, Dept of Information Technology, Pravara Rural Engineering College, Loni, India[1]
UG Student, Dept of Information Technology, Pravara Rural Engineering College, Loni, India[2]
UG Student, Dept of Information Technology, Pravara Rural Engineering College, Loni, India[3]

ABSTRACT: The purpose of image Steganalysis is to detect the presence of hidden messages in cover images. Steganalysis can be considered as a pattern recognition process to decide which class a test image belongs to: the cover image or the Stego image. In this paper, we present F5 algorithm for image Steganography and then collect the statistical evidence by using feature extraction and feature set value is classify by using support vector machine.. The propose work will give higher detection performance accuracy than comparative current Steganalysis. Therefore, we have introduced a SVM- classification methodology based on image statistical feature selection.

KEYWORDS: Steganalysis, Steganography ,Feature Extraction, Support vector machine classifier.

I. INTRODUCTION

The process of detecting hidden information inside the cover images is called as Steganalysis. The goal of Steganalysis is to collect sufficient evidence about the presence of embedded message[1]. Although images can be scanned for suspicious properties in a very basic ways, detecting hidden messages usually required a more technical approach, changes in size, file format, last modified time stamp, & in a color palette might point the existence of a hidden message.Security is important for keeping communications private. The two communicating parties don't want any third party to read, to view, to listen or to to manipulate confidential information. The information has to be unintelligible to others. In the present information based society, cryptography provides the much needed techniques for keeping information secret. In real world all communications can be intercepted by others also. Even when one uses any cryptographic technique to exchange information securely, that can be easily detected by intruders by seeing the encrypted text. One can do any kind of malicious things to the information so that the receiver is also not able to recover the original information. Therefore, there is a need for better secured communication mechanism other than cryptography to enable people to communicate securely. Because of this reason researchers have started research in making the communications invisible and with the help of information hiding techniques. The broad goal of a Steganalysis is to understand the effects of Steganographic data hiding into the cover medium. This knowledge is typically used to either strengthen the hiding systems or detect the use of Steganographic data hiding[1]. In order to develop any hiding scheme which is difficult to detect, it is necessary to analyse the resulting Stego objects. This is typically done by comparing statistical changes introduced while embedding the information. If a method causes distinct predictable changes in Stego objects it will be fairly easy to detect and modify the same. In this project the standard Steganographic method of least significant bit (LSB) insertion method is used. In support of the LSB insertion method we will used F5 algorithm for encryption of secret message for giving more security to message we can first compress the data then encrypt it[4]. After that we have to extract the features of image by using the MSD (Min Square

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 2, February 2014

Distortion) technique. After finding the feature extraction we use the SVM (Support Vector Machine) classifier to classify the image either it is a Stego image or a cover image which is the main significance of this project[2].

II. LITERATURE SURVEY

Many researchers are trying to discover statistical evidence of image steganalysis using svm as next generation discipline. Andreas Westfeld proposed the F5 algorithm hides data by modifying the quantized AC DCT coefficient indices[4]. Instead of overwriting the LSBs (like jsteg-jpeg), F5 decreases the absolute value of the indices, skipping indices with a value of 0. This preserves the symmetric and monotonic shape of the histogram. The approach was designed to discourage detection with the χ^2 test, a statistical test used to detect DCT coefficients that had been modified using jsteg-jpeg[4].

After embedding the message we extract the features. Bo Xu, Jiazhen Wang, Xiaqin Liu,Zhe Zhang work for extracting features we use image quality measures[5]. When we extract the features then it is easy to classify whether this image is cover image or Stego image. For classifying image we use Support Vector Machine(SVM)[2].Huan Dou, Zhipin Deng, Kebin Jia proposed supervised learning methods that it is a multi-classifier. We use Steganalysis methodology that consist of two phases. First is the Training phase where Multi Class SVM is trained for known class of images i.e with Clean images and Stego images embedded with Spatial, DCT and DWT domain embedding tools.

II. SYSTEM ARCHITECTURE

Proposed Steganalysis methodology consist of two phases .First is the Training phase where Multi Class SVM is trained for known class of images i.e. with Clean images and Stego images embedded with Spatial, DCT and DWT domain embedding tools. First the images are filtered with the Gaussian Low Pass Filter. Then a set of IQMs are calculated between clean or Stego images and their filtered version. These IQMs are gives as a features input to Multi class SVM, Based on these IQMs for known classes the Multi Class SVM determines the Trained Model. This trained model can be used for future testing for identifying the Stego images and also Stego images and also its hiding domain[2][3].

Second phase consist of Testing phase where images can be tested for hidden information .For this the test images are also filtered with same Gaussian Low Pass Filter and a set of IQMs is calculated between test images and its filtered version. These IQMs are given to Multi Class SVM for classification[5][6] .Based on trained model Multi Class SVM classifies the test images either as Stego or clean. The output of Multi class SVM is data hiding domain of test images, if it is Stego image.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 2, February 2014

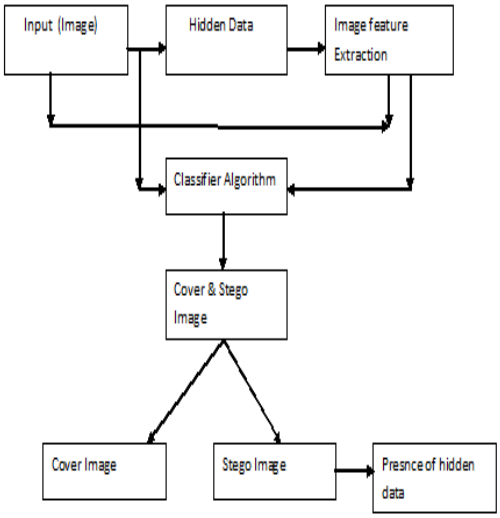

Figure 1.System Architecture

III. STEGANALYSIS

In recent years, digital Steganography has emerged as an increasingly active research area. Information can be hidden into images, Videos, and audios imperceptibly to human beings. They provides vast opportunities for covert communication using Steganographic techniques. On the other hand, Steganalysis can serve as an effective way to judge the security performance of Steganographic techniques. In other words a good Steganographic method should be imperceptible not only to human vision systems, but also statistically undetectable to the computer analysis.

Steganalysis refers to the body of techniques that are designed to detect and or estimate the hidden information from digital media with little or no knowledge of Steganography algorithm and its parameters[1].

A. Objectives

B. The main objectives of the proposed methodology are[1].

1. To find an efficient Steganalysis technique for images that take advantage of the statistical distortions present in it due to embedding with the help of image Quality Measures.
2. To classify a Stego image from cover image with high accuracy rate, using Support Vector Machine.
3. To give an overview where embedding algorithm developer can concentrate on particular distortions captured by Image Quality Measures to build efficient embedding algorithm.

IV. PROPOSED WORK

For secret message, we have used F5 algorithm. The F5 algorithm hides data by modifying the quantized AC DCT coefficient indices. Instead of overwriting the LSBs (like jsteg-jpeg), F5 decreases the absolute value of the indices, skipping indices with a value of 0[4]. This preserves the symmetric and monotonic shape of the histogram. The approach was designed to discourage detection with the χ^2 test, a statistical test used to detect DCT coefficients that had been modified using jsteg-jpeg[4].

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 2, February 2014

A. *Implementation Step Of F5 Algorithm.*
The algorithm F5 has the following coarse structure[4]:

1. Start JPEG compression. Stop after the quantization of coefficients.
2. Initialise a cryptographically strong random number generator with the key derived from the password.
3. Instantiate a permutation (two parameters: random generator and number of coefficients) .
4. Determine the parameter K from the capacity of the carrier medium, and the length of the secret message.
5. Calculate the code word length n = 2^k-1.
6. Embed the secret message with (1,n,k) matrix encoding.

(a) Fill a buffer with n nonzero coefficients.
(b) Hash this buffer (generate a hash value with k bit-places).
(c) Add the next k bits of message to the hash value (bit by bit, xor).
(d) If the sum is zero, the buffer is left unchanged otherwise the sum buffers index 1....n, the absolute value of its element has to be decremented.
(e) Test for shrinkage, i.e. whether we produce a zero. If so, adjust the buffer (eliminate the 0 by reading one more nonzero coefficient, i.e. repeat step 6(a) beginning from the same coefficient). If no shrinkage is occur advanced to new coefficients behind the actual buffer. If there is still message data continue with step 6a.

7. Continue JPEG compression (Huffman coding etc.).

B. *F5 Features.*

1. F5 has high embedding capacity (>13%) but can be pushed even further.
2. F5 has high embedding efficiency.
3. Resistance against both visual and statistical attacks.
4. Uses a common image format and carrier medium (JPEG).

C. *Image Quality Measures.*

Images quality measure (IQMs) are figure of merit used for the evaluation of imaging systems or of coding/processing techniques[5][6].
After embedding the message we extract the features. For extracting features we use image quality measures like-

a. Pixel differences based measures like Mean Square Distortion (MSD);
b. Correlation Based Measures, that is, correlation of pixel or of the vector angular directions;
c. Spectral distances based measures, that is , Fourier Magnitude or Phase Spectral discrepancy on a block basis;
d. Edge based measures, that is, displacement of edge positions or their consistency across resolution levels.

D. *Support Vector Machine (SVM) Classifier*

When we extract the features then it is easy to classify whether this image is cover image or Stego image. For classifying image we use Support Vector Machine (SVM)[2][7]. It is a multi-classifier. We use Steganalysis methodology that consists of two phases. First is the Training phase where Multi Class SVM is trained for known class of images i.e. with Clean images and Stego images embedded with Spatial, DCT and DWT domain embedding tools. First the images are filtered with the Gaussian Low Pass Filter. Then a set of IQMs are calculated between Clean or Stego images and their filtered version. These IQMs are gives as a features input to Multi class SVM, Based on these IQMs for known classes the Multi Class SVM determines the Trained Model[7]. This trained model can be used for future testing for identifying the Stego images and also Stego images and also its hiding domain. Second phase consists

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 2, February 2014

of Testing phase where images can be tested for hidden information. For this the test images are also filtered with same Gaussian Low Pass Filter and a set of IQMs[3] are calculated between test images and its filtered version. Based on trained model Multi Class SVM classifies the test images either as stego or clean[2][5].

II.SVM Structure.

Following points shows support vector machine structure are as follows[2].

1)Firstly the encoding information are extracted from reference frames and encoded MBs frames from current frame.

2)The first stage SVM classifier is used to identify the MBs in homogeneous areas, which are usually coded with SKIP mode.

3) If the identified MB mode(classified by the first stage) is not SKIP mode, the second stage SVM mode, the current MB is encoding by SKIP mode and Inter 16x16 mode.

4) If the identified MB mode (classified by the first stage) is not SKIP mode, the second stage SVM classifier is used to identify the MBs in big-partition modes, which are usually coded with Inter 16x16 mode.

5) If the identified MB mode (classified by the second stage) is Inter16x16 mode, the current MB is encoded by SKIP mode and Inter 16*16 mode (to ensure the encoding quality);

6) If the identified MB mode (classified by the second stage) is not Inter 16*16 mode, the third stage SVM classifier is used to identify the MBs in comparatively big partition modes, which are usually coded with Inter 16*8 mode or Inter 8*16 mode;

7) If the identified MB mode (classified the third stage) is Inter 16*16 mode, the current MB is encoded by Inter 16*8 mode and Inter 8*16 mode, and to ensure the encoding quality Inter 16*16 mode is also used to encode;

8) If the identified MB mode (classified by the third stage) is not Inter 16*16 mode, the current MB is encoded by small-partition modes, which are Inter 8*8 mode, Intra 16 mode, Intra 8 mode and Intra 4 mode;

9) Choose the optimal mode by rate-distortion performance method and then encode the next MB.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 2, February 2014

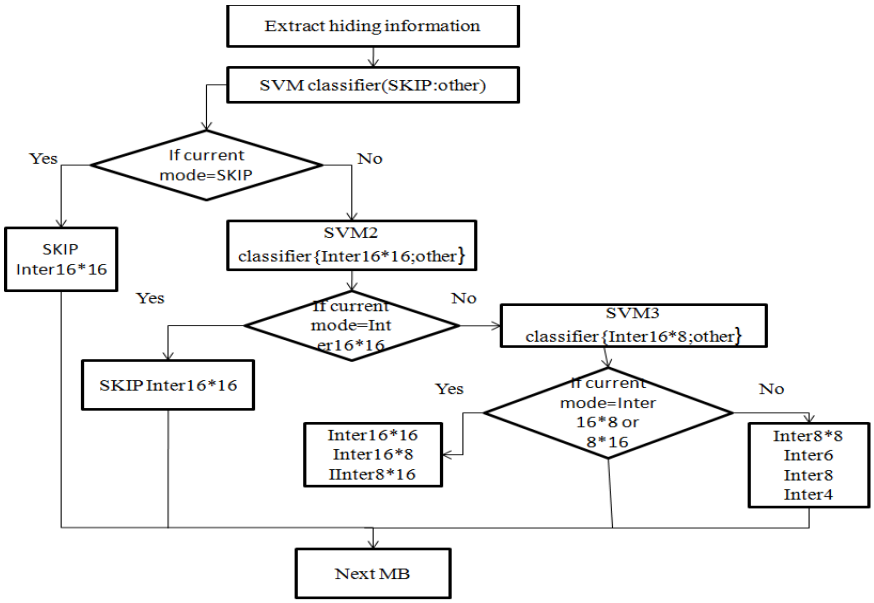

Figure 3. SVM Structure

V. CONCLUSION

In this paper, we have addressed the problem of Steganalysis of images, and we have developed a Stegnalysis Tool for discriminating between cover images and Stego images. Our approach is based on the hypothesis that message embedding schemes leave statistical evidence or structure in images that can be exploited for detection. After selecting an appropriate feature set, we used multi class support vector machine as a classifier for classifying Stego images from non Stego images.

ACKNOWLEDGMENT

we would like to express our sincere appreciations to our Project Guide **Prof.Mr.G.R.Suryawanshi** from the Information Technology Department whose help, stimulating suggestions and encouragement helped us all the time to complete this paper work, also many thanks thanks to **Prof.Mrs.V.V.Mandhare** our project Co-Ordinator Department of Information Technology for giving us timely suggestions to commence this project in the first instance and necessary research work.

REFERENCES

[1] J.R. Krenn, "Steganography and Steganalysis", January 2004.
[2] Huan Dou, Zhipin Deng, Kebin Jia, "A Fast Macroblock Mode Decision Algorithm for MVC Based on SVM", Dept. of Electronic Information & Control Engineering, Beijing University of Technology, Beijing, China.
[3] I. Avcibas, N. Memon, and B. Sankur. "Steganalysis using image quality metrices", Image Processing, IEEE Transactions, VOL. 12, NO. 2, FEBRUARY 2003.
[4] Andreas Westfeld, "F5—A Steganographic Algorithm High Capacity Despite Better Steganalysis", Technische University at Dresden, Institute for System Architecture Dresden, Germany.
[5] Bo Xu, Jiazhen Wang, Xiaqin Liu,Zhe Zhang. "Passive Steganalysis Using Image Quality Metrics and Multi-class Support Vector Machine," Third International Conference on Natural Compution(ICNC 2007)IEEE.
[6] I. Avcibas, B. SAnkur. "Statistical evaluation of image quality measures," Journal of Electronic Imaging , Vol. 11(2), April 2002.
[7] Tristan Fletcher, "Support Vector Machines Explained".