Steganalysis with Location Based Features and Efficiency

Mrs. K. Rajasri, M. Tech.,

Senior Assistant Professor,

Christ College of Engineering and Technology, Puducherry.

Ms. T. Indhumathi, M.Tech.,

Student,

Christ College of Engineering and Technology, Puducherry.

Abstract

Data hiding is a technique of concealing top secret information into a cover media and preventing an observer from being aware of the subsistence of the hidden communication. According to the troubles of steganography, the main endeavour is to provide a better imperceptibility of stego-image that can be prepared by diminishing distortion of image. One of the popular techniques in data hiding is steganography in which the simple method for image hiding is the Least Significant Bit substitution method. There are two image hiding techniques to improve the quality of the stego-image. The first one is to discover the best block matching matrix and the other one is to find the optimal substitution matrix. There are two various approaches for optimal substitution matrix. The first method is the global optimal substitution and the second one is local optimal substitution. If we transform each number in the pixel value into a number divisible by 5, then this will not affect the Human Visual System. Then, the pixel value of secret image is transformed to new pixel value.

1. Introduction

The word steganography is of the Greek origin and translates as "covered writing". Steganography is a dynamic means with a long history and the potential to get used to the new levels of technology. Steganography is the method of hiding secret or sensitive communication within some medium that appears to be the usual message transfer. Apart from the sender and the receiver, no one knows the presence of the message. This is to protect the data from illegal or unnecessary viewing. Steganography has evolved into a digital plan of hiding information in some form of media, such as an image, an audio file and video file or even in Transmission Control Protocol header. Steganography be a technique for secure communication, the stego-images do not consist of any evident artifacts due to message embedding. In other language, the set of stego-images should have the unchanged arithmetic properties like the set of cover-images. If there is an algorithm that can assess whether or not a given image contains secret information with an achievement rate better than random guessing, the steganographic method is considered broken down.

Steganography may provoke negative effects in the view of individual confidentiality, business action, and nationwide safety measures. The criminals can misuse this practice for planning illegal actions. Terrorists may possibly also use these techniques to assist for international attacks and stop themselves from being traced. Some others may still feel the chance of transmitting the computer virus or Trojan horses via data concealing methods. In general, steganography may be used as a tool to cover up proof of any criminal action and consequently make difficult the work of forensic computer analysts trying to decipher exactly what data resides on a suspect's computer hard drive. Thus, it raises the concerns of enhancing wardens' potential and cutback these harmful effects by developing the methods of "steganalysis". Most steganographic systems nowadays hide communication by slightly modifying an existing cover object, such as a digital image.

Steganalysis involves detecting the employ of steganography in a file with little or no information about the steganography algorithm and/or its parameters. Steganographic algorithms occasionally leave a mark in the file that is encrypted. With this information, the presence of secret message can be detected. It is reasonable to reveal that steganalysis is both an art and a science. The ability of steganalysis plays a major function in the selection of features or quality to test for hidden communication while the science helps in designing the tests themselves. Many steganalytic techniques which have been emerged in recent times might come under one of these two classes:

ad hoc schemes and feature based schemes that are general and that employ classifiers to distinguish original and stego samples. Feature based steganalysis is the up-and-coming division in information forensics and security. Its final aspire is to recognize the existence of a clandestine message.

2. Literature Survey

Steganography is an antique art but digital knowledge gives it novel way so that can be hide information in digital images and signals also. The goal of steganography is to insert a message within an innocuous looking cover medium so that casual assessment of the resulting medium will not disclose the existence of the message.

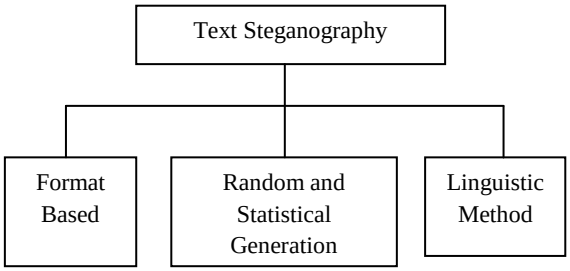

Figure 1.Three basic categories of text steganography [1]

When a new steganographic method is proposed, it is necessary that it not be noticeable using known feature sets. [3] A stego-key is used to manage the hiding procedure so as to limit uncovering and rescue of the secret communication. [10]

In LSB methods, secret communication can be embedded by replacing unimportant or redundant parts of a cover image. Mainly, in LSB method of Steganography, bits of secret message are substituted into the least significant bits of cover image's pixel. The major advantages of substitution method are simple to employ and its ease. But, LSB techniques are highly susceptible because slight modifications in least significant bits of cover image can destroy entire communication.

The Perceptron is an instance of an easy linear algorithm that is very fast. [7] Simple LSB techniques are easier to execute but produces low quality Stego images. So, to overcome this difficulty new technique is proposed. This technique is based on a genetic algorithm to embed secret information inside host image. But, this algorithm requires large computation time for approximate solutions.

There are various LSB detectors like RS, SPA, and DIH etc. Dynamic approach is proposed to decrease computational time. This approach selects best answer from all possible solutions. In 2010, a new scheme called Transforming LSB substitution technique is proposed to overcome problem linked with above two approaches.

Steganalysis is the capability and discipline of exposure of the existence of steganography. Steganalysis is a means of detecting clandestine message hidden using steganography. The goal of steganalysis is to gather sufficient evidence about the existence of embedded message and to break the security of its carrier. Thus break the security provided by steganography. Both steganography and steganalysis have received a lot of notice from law enforcement and media. The battle stuck between steganography and steganalysis is never comes into conclusion. The final goal of steganalysis is to make a decision if an image contains an embedded message. As this field has developed, determining the piece of the message and the actual contents of the message are also becoming active areas of research.

To convert any colour to a gray scale image of its luminance, first one must obtain the values of its red, green, and blue primaries in linear intensity encoding, by gamma expansion. The performance of detectors built using machine learning tools may be quite dissimilar from the clairvoyant Classifier. [5] However, in steganalysis, the main content of an image is not an issue to be considered since human eyes cannot tell the dissimilarity between an original image and its stego-version. [4] Feature selection, as a pre processing step to machine learning, is effective in reducing dimensionality, removing unrelated information, increasing learning accuracy and improving result unambiguousness. [8]

First, a set of statistics called steganalysis characteristics is extracted from a pair of training set which contains cover and stego mediums correspondingly. [9] Following feature selection in the context of steganalysis provides numerous merits. [6]

- Irrelevant features are removed.
- A more rational method can be followed for classifier based steganalysis.
- The classification performance is improved.

There are two approaches to the crisis of steganalysis. One is to come up with a steganalysis means exact for a particular steganographic algorithm. The other is just beginning techniques which are detached of the steganographic algorithm to be considered. Each of the two methods has its own merits and demerits.

The importance of steganalytic techniques that can reliably notice the existence of concealed information in images is rising. Steganalysis is classified into: Statistical Steganalysis and Signature Steganalysis. When secret data hides in an image then the statistics of an image altered. Due to add of this secret information in the image, its pixel values vary. This alter in statistic of the image is used during investigation to notice the secret data.

3. Existing System

In cryptography, the data is noticeable but not in any evocative variety. Only by knowing the cryptographic algorithm, the hidden data can be deciphered**.** In cryptography, every person knows that there is hidden information present. But only the right algorithm can reveal. I.e. in cryptography a communication can be easily seen and recognized as cryptic message. But only the person who has information as how the data is encrypted will come to know how to decrypt it. In cryptography plaintext/secret message is encoded into cipher form. So, an attacker cannot easily decode cipher text into original secret communication. On the other hand, Steganography is used to hide the presence of the top secret payload. The original dataset was split into 3 categories, nevertheless only two of these are used.

Whereas the data in Steganography is written in plain text but is hidden in the cover medium so that it's tough to notice and uses the non-prominent area of the text or image or video i.e. any medium which is being used as a Steganographic cover medium, it has various examples. But a simplest can be a picture of a man has a pose in which his figure points up means he is glad and downward points mean he is miserable. Therefore the sender and receiver know the code and can identify the hidden information. To hide the information within any media involve vital characteristics like a cover medium or file which is necessary for hiding data , a secret data that required to be concealed and a key or code word that may be used by sender and recipient for encryption and decryption. In brief, steganography can be signified as

Secret data + cover medium = stegogramme

Stegogramme + stego key = stego-medium.

Network steganography shelters a broad range of techniques. The typical steganographic method utilizes digitized media files as a cover medium for hiding data. Network steganography uses communication protocols such as TCP/IP. The means to detect the hidden information in communication is called as steganalysis.

Pictures are the most frequent and convenient way of conveying or transmitting information. An image is worth a thousand terms. Pictures concisely express information about positions, sizes and inter relationships between objects. They represent spatial information that we can identify as objects. Human beings are excellent at deriving information from such images, because of our innate visual and mental abilities. About 75% of the information received by human is in pictorial appearance. A picture is digitized to renovate it to an appearance which can be stored in a computer's memory or on some look of storage media such as a hard disk or compact disk. This digitization method can be completed by a scanner, or by a video camera associated to a frame grabber panel in a PC. Once the image has been digitized, it can be operated upon by different image processing functions.

Image processing operations can be approximately distinguished into three foremost categories. They are: Image Compression, Image Enhancement and Restoration, and Measurement Extraction. It involves dropping the amount of memory required to store a digital image. Image defects can be caused by the digitization process or by faults in the imaging set-up. These can be corrected using the Image Enhancement methods. Once the image is in good form, the Measurement Extraction operations can be used to obtain useful features from the image. Each pixel in the image is stored as a number between 0 and 255, where 0 represents a black pixel, 255 represents a white pixel and values in-between represent shades of grey.

In packet length based Steganography, the piece of the transmitted packet is being modified to conceal the data and analysing various data packets, and it's feasibility to detect the data present in it. The detector has to study a large amount of packets to detect the inconsistency or hidden data.

3.1 Data Hiding By Pixel Mapping Method

In 2010, a new technique is projected to plot information into image called pixel mapping method. It uses idea of pixel intensity and number of one's in pixel to map information. This method produces improved embedding ability and PSNR Value over PVD.

Now, steganography of the new era is growing with splendidly greater opportunity for harm. Through the latest technological advancements, the restriction on the extent of the secret message has been removed. Consider instance involving the use of Skype where it requires a carrier which can be an MP3 song or a video—there was no such requirement for the transmission of a photograph. The data were hidden in the bits of a digital VOIP conversation. In this fresh age of steganography, the scapegoat that co-conspirators are using is not the carrier but the whole communication protocol with an advantage longer the communicators talk, the longer can be the secret message which is sent. It makes the data nearly impossible to detect. In 2011, PMM method is proposed with BPCS which produces better image quality over PMM method.

4. Proposed System

This proposed scheme is to construct a convenient steganographic implementation to hide content inside grey scale images. The secret communication is hidden within the cover image using Five Modulus technique. The novel algorithm is called FMT.

4.1 Five Modulus Technique

FMT consists of transforming all the pixels in the 5X5 window size into its corresponding multiples of 5. After that, the secret message is hidden in the 5X5 window as non-multiples of 5. Since the modulus of non-multiples of 5 is 1, 2, 3 and 4, therefore; if the reminder is one of these, then this pixel indicates a secret message. The secret key that has to be sent is the window size.

The primary idea behind FMT is based upon the following idea: A common characteristic in most of images is that the neighbouring pixels are interrelated. So, for bi-level images, the neighbours of a pixel have a tendency to be similar to the original pixel. Therefore, FMT consists of dividing the image into blocks of k X k pixels each. Clearly, in bi-level grey images, we know that each pixel is a numeral between 0 and 255. Hence, if we can transform each number in that range into a digit divisible by 5, then this will not affect the Human Visual System. The basic idea in FMT is to verify the whole pixels in the k X k block and transform each pixel into a number divisible by 5 according to the algorithm.

For extracting the required secret message from the stego image, the location based features are considered. While considering the features of an image based on the location of the object and its neighbourhood, the efficiency of the steganalysis process is improved. This proposed approach not only provides larger embedding ability but also results in an acceptable Stego image quality that can be seen by human eyes. The main benefit of this novel method is to keep the dimension of the cover image stable while the secret message increased in dimension. PSNR is captured for each of the images tested. Based on the PSNR estimate of each images, the stego image has peak PSNR value. Hence this new steganography method is extremely proficient to hide the data inside the image.

5. Conclusion

Though LSB embedding methods hide data in such a way that human does not perceive it, these embeddings often can be easily destroyed by compression, filtering or a less than perfect format or size conversion. Hence, it is often necessary to employ sophisticated techniques to improve embedding reliability. This is a steganalytic method based on integrating singular values calculated over image sub-blocks resulting in a new robust steganalyzer based on the features in the image calculated along with the location details.

6. References

[1] 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.

[2] Sonam Chhikara, Parvinder Singh, "SBHCS: Spike based Histogram Comparison Steganalysis Technique", International Journal of Computer Applications (0975 – 8887) Volume 75– No.5, August 2013.

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

[4] Yun Q. Shi, Guorong Xuan, Dekun Zou and Jianjiong Gao, "Image Steganalysis Based on Moments of Characteristic Functions Using Wavelet Decomposition, Prediction-Error Image, and Neural Network".

[5] Jan Kodovský and Jessica Fridrich, "Steganalysis in high dimensions: Fusing classifiers built on random subspaces".

[6] S.Geetha and Dr.N.Kamaraj, "Optimized Image Steganalysis Through Feature Selection Using MBEGA", International Journal of Computer Networks & Communications (IJCNC), Vol.2, No.4, July 2010.

[7] Ivans Lubenko and Andrew D. Ker,"Going from Small to Large Data in Steganalysis".

[8] Lei Yu and Huan Liu, "Feature Selection for High-Dimensional Data: A Fast Correlation-Based Filter Solution".

[9] B. B. Xia, X. F. Zhao and D. G. Feng, "Improve Steganalysis by MWM Feature Selection".

[10] Firas A. Jassim, "A Novel Steganography Algorithm for Hiding Text in Image using Five Modulus Method", International Journal of Computer Applications (0975 – 8887) Volume 72– No.17, June 2013.