International Journal of Emerging Trends & Technology in Computer Science (IJETTCS)
Web Site: www.ijettcs.org Email: editor@ijettcs.org
Volume 4, Issue 1, January-February 2015 **ISSN 2278-6856**

A Survey on Digital Image Steganography

Kalaivanan.S[1], Ananth.V[2] and Manikandan.T[3]

School of Engineering and Technology, Department of Computer Science, Pondicherry University, Puducherry, India

Abstract

This paper describes the various developments in the field of Image Steganography. It includes applications and the different techniques used to implement the various algorithms in use whilst comparing them on the basis of a few parameters integral to the field of Image Steganography

Index Terms:-Image, Steganography, DCT, DWT, DFT, Adaptive, Model Based, LSB, Spatial, Frequency, JSteg, F5, ABCDE, Distributed

1.Introduction

Steganography is a technique to hide information in ways that prevent the detection of hidden messages. It uses digital media as carriers for secret communication. Cryptography and Steganography are not one and the same. While Cryptography scrambles a message so that it cannot be understood, Steganography hides the messages so that it cannot be seen. Un-detectability, Robustness and capacity of the hidden data are the main features that differentiate steganography from cryptography. A Steganogram is nothing but a steganographically modified carrier with hidden information. Secure steganographic algorithms hide confidential messages in carrier media to form steganograms so that an attacker will not able to find it. In this paper we are going to touch upon image steganography where image file is used as a carrier to make steganograms, or stego-images.

2.Literature Survey

Information hiding techniques have become an important research area in recent years, ever since researchers realized that developing techniques to solve unauthorized copying, tampering, and distribution of multimedia data via Internet was urgent. The information hiding techniques include convert communication, steganography, and digital watermarking. Steganography employs an innocent-looking media called a host image to imperceptibly carry secret data to an intended recipient. The image embedded with the secret data is called a stego-image, or cover image, and looks like a normal image. Unintended recipients of a stego-image are unaware of the existence of the hidden data. The camouflage of secret data in stego-images to pass sensitive data between two parties can be regarded as secret communication, but it is different from

cryptography

Generally, steganographic methods proposed in the past few years can be categorized into two types. The methods of the first type employ the spatial domain of a host image to hide secret data. In other words, secret data are directly embedded into the pixels of the host image[1,2,3,4]. Steganographic methods of the second type employ the transformed domain of a host image to hide secret data[5,6,7].Transformation functions like the discrete cosine transform (DCT) or discrete wavelet transform (DWT) are first exploited to transform the pixel values in the spatial domain to coefficients in the frequency domain. Then the secret data are embedded in the coefficients. Most of previous works on steganography aimed at hiding grayscale images in grayscale host images and excluded the use of color images because of the complexity of using color images as host images. This constraint has two drawbacks. One is that only grayscale images are used to hide grayscale images. This is impractical because many valuable images are available in color. The other is the limited hiding capacity that a grayscale host image has.

The scheme proposed in [8] hid large amounts of data in the pixels of a true color image, which used 8 bits to represent each color component of a color pixel. The secret image can be both a grayscale and a true color image. However, the quality of the extracted color secret image is not good in terms of the peak signal-to-noise ratio(PSNR) value and visual observation.

3.steganography applications

This section describes some of the well known apllications of steaganography. Steganography is utilized in different valuable applications,e.g., copyright control of materials, upgrading strength of picture web crawlers and smart IDs (character cards) where people's information are implanted in their photos. Steganography would give an extreme assurance of confirmation that no other security device may guarantee. Motivated by the thought that steganography can be inserted as a component of the typical printing procedure. Different applications are video–audio synchronization, organizations' sheltered dissemination of mystery information, TV, TCP/IP packets[9]furthermore checksum embedding[13]. Steganography is appropriate to, yet not constrained to confidential correspondence and mystery information storing,Media Database systems,Access control framework for advanced substance appropriation, Protecting alteration of the data.

4.steganography techniques

Some of the major techniques used in the field of Image Steganography are mentioned below. Some trivial algorithms utilizing the techniques are also listed.

A.Exploiting Image Format

These techniques exploit the format of the digital image to hide secret data in areas that are not used for storing data. These techniques are robust to image manipulation as the image data is left untouched. The data is not well hidden and may be revealed accidentally to unintended users.

1.Exploiting EOF symbol

This technique utilizes the EOF character that is used to represent a condition in which no more data can be read from the data source. The EOF character is digital images by appending data such as text after the end of file. Files can be viewed as an image file in image viewer. The textual data can be seen when viewed in a text viewer[9]. This technique enables us to hide a large payload but is not robust to image modifications.

2.Exploiting EXIF Format

In this technique we use the EXIF data used in digital image to store some data. EXIF or Exchangeable Image File Format is used to describe the camera settings in which the picture was taken. Secret data can be inserted in the comments field in the EXIF data of the image file as it often requires special software to require EXIF data. When the image is opened in image viewer the textual data is not displayed[9]. The payload is sufficiently large as EXIF specification allows data up to 64KB in size. Also the hidden data cannot be easily destroyed if subjected to image manipulation as EXIF data is retained unless specifically intended to do so.

B.Spatial Domain Techniques

Spatial Domain Techniques consists of techniques which modify the pixels of the image in spatial domain in order to hide the message inside it.

1.LSB Insertion

In LSB insertion we embed the secret data in the LSBs of the pixels to take the advantage of inability of Human Visual System to detect minute changes in intensity or colour of the image. The secret data is converted into a bit-stream and then the LSBs of the pixels of the image are replaced by the bit by bit for up to a maximum of 4 bits of a pixel's colorspace[10][11]. This scheme allows sufficient payload size to data size ratio but is not immune to image modification. Also it is easily detected by statistical attacks. The cover image must be carefully selected so that there are no large areas of same colour and intensity as embedding in such images will allow the Human Vision System to detect the changes incurred by LSB embedding. Also it requires lossless image formats such as bmp, tiff or RAW.

3.Histogram Based Data Hiding

In this class of algorithms we try to distribute the secret data amongst the pixels whose intensity has the highest frequency. This increases its robustness to some forms of histogram based statistical attacks[13]. HKC Algorithm: In this technique the shift the intensities of the pixels that fall between the intensity values of the peak point and zero point of the histogram of the image and embed the secret data in the pixels that fall in the peak point intensity[12].

C.Frequency Domain Techniques

1.Discrete Cosine Transform

Discrete Cosine Transformation is useful in storing secret data as DCT is used to compress the data as quantized frequency coefficients in video and image compression extensively. The description of 2D-DCT for an input image F and an output image T is calculated as

$$T_{pq} = \alpha_p \alpha_q \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} F_{mn} \cos\frac{\pi(2m+1)p}{2M} \cos\frac{\pi(2n+1)q}{2N}$$

Algorithm 2: F5
Input: message, shared secret, cover image
Output: stego image
initialize PRNG with shared secret
permutate DCT coefficients with PRNG
determine k from image capacity
calculate code word length n←2k– 1
while data left to embed do
get next k-bit message block
repeat
G←{n non-zero AC coefficients}
s←k-bit hash fof LSB in G
s←s⊕ k-bit message block
if s≠0 then
decrement absolute value of DCT coefficient Gs
insert Gs into stego image
end if
until s= 0 or Gs≠0
insert DCT coefficients from G into stego-image

where

$$0 \le p \le M-1$$
$$0 \le q \le N-1$$

and M, N are the dimensions of the input image while m, n are variables ranging from 0 to M-1 and 0 to N-1respectively JSteg Algorithm: This algorithm created by Derek Upham replaces the LSB of the frequency coefficient by the secret message. The embedding mechanism skips all DCT coefficients with the values 0 or 1. However, it disturbs the bell curve of the DCT coefficients of JPEG compression and as a result it can be easily detected by $\chi 2$ (chi square) test. The JSteg algorithm is outlined in algorithm 1.

Algorithm 1:JSteg
Input: message, cover image
Output: stego-image
while data left to embed do
get next DCT coefficient from cover image
if DCT ≠0 and DCT ≠1 then
get next LSB from message
replace DCT LSB with message LSB
end if
insert DCT into stego image
end while

F5 Algorithm: This algorithm created by Andreas WestField embeds the secret data into non-zero AC DCT coefficients by decreasing the absolute value of the coefficient by 1. It largely remained undetected from $\chi 2$ test and their variations but was soon detected by Fridrich Algorithm. The F5 algorithm is outlined in algorithm 2.

2.Discrete Wavelet Transform

JPEG2000 allows use of wavelets for compression in place of DCT. Embedding in DWT domain tends to be more flexible and outperforms DCT in terms of compression survival. But the limited capacity of DWT methods makes it less interesting[9].

3.Discrete Fourier Transform

Discrete Fourier Transform can also be used to embed secret data into the frequency domain, but the inverse discrete Fourier (iDFT) encompasses round-off error when converting real values to integer values which makes it difficult to obtain the hidden message accurately[14].

D.Adaptive Steganography

Adaptive Steganography is a special form of the above methods. Also known as "Statistics aware embedding", this method takes into consideration statistical global features of the image before attempting to interact with its LSB/DCT coefficients.

1.Model Based Method(MB1)

In this algorithm the stego-image is generated using a generalized Cauchy distribution that results in minimum distortion. However it can be easily detected by analysing the differences of "blockyness" between a stego-image and its estimated original image[9].

2.A Block Complexity based Data Embedding (ABCDE)

This algorithm by Hioki utilises the Edge Embedding technique to embed the secret data in the cover image. Edge Embedding follows edge segment locations in a fixed block with its centre on an edge pixel. Pixel data of noisy blocks are replaced with another noisy block with the embedded data. Suitability of replacement blocks is determined by run length irregularity and border noisiness. Manual configuration of certain important control parameters renders the algorithm useless for automatic processes.

E.Distributed Image Steganography

This technique utilizes the (k, n) Secret Sharing Scheme proposed by Adi Shamir to create k shares of the data and then embeds it into LSB or Frequency Domain Coefficients of n images. Atleast k stego-images are required to extract the secret data successfully.

5.ANALYSIS

A.Parameters in Consideration

1. **Type of Cover Image:** It is very important that the technique is able to be implemented to be applied to a range of image formats as frequent use of any particular format or type of image will arouse suspicion.
2. **Payload Size:** Delivery of the payload without detection is the main aim of image steganography and hence the algorithm must be able to hide the entire message in the cover media available.
3. **Robustness:** The algorithm must hide the secret data in such a way that minor image processing should not destroy the hidden data altogether. Also ideally it should be able to handle minor errors and correct it to extract the secret message
4. **Detectability:**

a.Human Vision System: The output generated from the algorithm must be visually similar to the human eye.

Statistical Attacks: Statistical Attacks such as $\chi 2$ test, StegDetect and Fridrich Algorithms can detect the presence of hidden data by analyzing the composition of the image by statistical means. Once suspected the image file may be subjected to further analysis to find the encoding algorithm hence revealing the secret data.The comparison of the various techniques in terms of generic steganographic parameters are given in Table 1.

6.TABLES AND FIGURES

TABLE I. COMPARISON OF VARIOUSTECHNIQUES

Parameters	Format exploit	LSB	DCT	DWT	DFT	Adaptive	Distributed
Cover Format	Any	Any	JPG	JPG	BMP, TIFF	BMP,TIFF	JPG,TIFF, BMP
Robustness to modification	Good	Bad	Good	Good	Good	Good	Bad
Payload size	Large	Large	Medium	Small	Small	Tiny	Large
Visual Detection (PSNR)	∞	High	Medium	Low	Low	Low	High
Steganalysis	EXIF viewer, Text editor	Spectral Analysis	X^2 test	RS Analysis	Unknown	Spectral Analysis	X^2 test

7. CONCLUSION

This paper discussed the various major algorithms of each type in the domain of Image Steganography. It also presented a table for comparing the various pro and cons of each type of Steganography Algorithms so that it is easy to determine the trade-offs to be negotiated for improving a particular attribute. Nowadays Frequency Domain techniques are the main focus of research as Spatial Domain techniques have all been detected by Steganalysis tools. But Frequency Domain techniques require more processing power to embed information as

well as during extraction; certain transformations may require a few minutes for a small image. As such mass surveillance of Frequency Domain techniques is not feasible and hence are an excellent candidate for Image Steganography.

REFERENCES

[1] C.K. Chan, L.M. Cheng, Hiding data in images by simple LSB substitution, Pattern Recognition 37 (2004) 469–474.

[2] C.C. Chang, J.Y. Hsiao, C.S. Chan, Finding optimal least-significant bit substitution in image hiding by dynamic programming strategy, Pattern Recognition 36 (2003) 1583–1595.

[3] Chih-Ching Thien, Ja-Chen Lin, A simple and high-hiding capacity method for hiding digit-by-digit data in images based on modulus function, Pattern Recognition 36 (2003) 2875–2881.

[4] R.Z. Wang, C.F. Lin, J.C. Lin, Image hiding by optimal LSB substitution and genetic algorithm, Pattern Recognition 34 (2001) 671–683.

[5] C.C. Chang, T.S. Chen, L.Z. Chung, A steganographic method based upon JPEG and quantization table modification, Information Sciences 141 (2002) 123–138.

[6] M. Iwata, K. Miyake, A. Shiozaki, Digital steganography utilizing features of JPEG images, IEICE Transactions on Fundamentals E87- A (2004) 929–936.

[7] H. Noda, J. Spaulding, M.N. Shirazi, E. Kawaguchi, Application of bit-plane decomposition steganography to JPEG2000 encoded images, IEEE Signal Processing Letters 9 (2002) 410–413.

[8] M.H. Lin, Y.C. Hu, C.C. Chang, Both color and gray scale secret images hiding in a color image, International Journal of Pattern Recognition and Artificial Intelligence 16 (2002) 697–713.)

[9] N.F. Johnson, S. Jajodia, Exploring Steganography: Seeing The Unseen, IEEE Computer 31 (2) (1998) 26–34.

[10] V.M. Potdar, S. Han, E. Chang, Fingerprinted Secret Sharing Steganography For Robustness Against Image Cropping Attacks, In: Proceedings Of Ieee Third International Conference On Industrial Informatics (Indin), Perth, Australia, 10–12 August 2005, pp. 717–724.

[11] J.H. Hwang, J.W. Kim, And J.U. Choi, 'A Reversible Watermarking Based On Histogram Shifting,' IWDW 2006, LNCS 4283, pp.348-361, 2006.

[12] Wen-Chung Kuo, Dong-Jin Jiang and Yu-Chih Huang, 'Reversible Data Hiding Based on Histogram', Advanced Intelligent Computing Theories and Applications with Aspects of Artificial Intelligence Lecture Notes in Computer Science, Volume 4682, 2007, pp 1152-1161.

[13] W. Bender, W. Butera, D. Gruhl, R. Hwang, F.J. Paiz, S. Pogreb,Applications for data hiding, IBM Systems Journal 39 (3&4)(2000)547–5.