Survey on Different Data Hiding Methods For Binary Host Images

Sofia. A. Khan[1], Antara Bhattacharya[2]

Student, MTech, Computer Science and Engineering, GHRIETW, Nagpur, India

Assistant Professor, Computer Science and Engineering, GHRIETW, Nagpur, India

Abstract: ***Data hiding approach aims to embed some secret information into a carrier data signal by altering the unimportant components for copyright protection or secret communication. In some cases, the data-hiding operation will lead to distortion in the host image. However, this distortion is unacceptable to some applications, e.g., medical or military images. In this case it is critical to embed the supplementary secret message with a reversible manner so that the original contents can be perfectly restored after extraction of the hidden data. There are number of information hiding techniques with the features that not only the message needs to be exactly extracted, but also the cover itself should be restored losslessly. This characteristic is important in some special context such as law forensics, military imagery and medical imagery. In these applications, the cover is very valuable or important to be damaged. So in the following text various techniques for data hiding are discussed. The survey is about how the data hiding techniques have gradually evolved with time and how techniques and methods when merged together can be effective.***

Keywords: data hiding; cryptography; secret message; encryption; decryption

1. Introduction

The expanding popularity of digital media has accompanied in concern over security-related issues. The privateness of a document is typically achieved by encryption. However, as an encrypted message normally discloses the importance of its content, the cipher text also attracts the interest of cryptanalysts. Steganography differs from encryption, in that it embeds censorious information in a noncensorious host message (e.g., webpages and advertisements) to distract opponents' attention.

Steganography is a study of techniques that embed secret information unnoticeably into a cover medium for security protection or covert communication purposes. Mostly the research is concentrated on images, audios, and videos as cover media. Imperceptibility of data hiding is commonly achieved by utilizing the weaknesses of the visual systems and human auditory, using the techniques for shifting lines, words, or characters by a small amount in an image containing text. Other techniques hide data by adding inessential data, or making use of alternative representations of electronic data. Such as, various combinations of the color palette entries in a GIF image can be used to hide data in the image

Data hiding, parallel to invisible watermarking and steganography, is the art of hiding a secret message into an innocent-looking host media by sustaining the least perceptional distortions. With the fast escalation of digital multimedia, data hiding has become useful to secure digital contents from illegal distribution and from spiteful tampering. A binary image requires only 1 bit per pixel as compared with 8 bits per gray pixel or 24 bits per color pixel. The small memory or storage requirement makes a binary image an ideal format for digitizing, processing, transmitting and filing large amount of daily documents whose contents are typically black and white in nature.

The basic idea is to have a cryptographic image computed based on the region , together with some additional data, e.g., the owner's information, is embedded by altering the region to generate the secure image[1]. The image visually resembles the host image with very few perceptible distortions. Upon verification, the image is divided into the identical regions. The secret message is then extracted from host image, decrypted and compared with the recomputed image for image integrity verification.

2. Related Work

a) Yu-Chee Tseng,Yu-Yuan Chen, and Hsiang-Kuang Pan proposed in, "A secure data hiding scheme for binary images", enhancement in the image hiding quality and hiding capacity, author presented a novel scheme which can hide a some amount of data by changing number of bits in the original binary image using secret keys as binary matrix and an integer weight matrix. The operator XOR is used so that the keys are not compromised. Another function of the weight matrix is to intensify the data-hiding capacity.

b) S.-C. Pei and J.-M. Guo proposed in, "Hybrid pixel-based data hiding and Blockbased Watermarking for error-diffused halftone images" that digital halftoning is a technique for improving graylevel images into two-tone binary images. These images can favour the original images when observing from a distance by the human visual system. Techniques in the first classification embeds invisible binary data into halftone images, which further can be recovered by scanning and applying some extraction algorithms. Methods in the second category embed hidden visual patterns into two or many halftone images so that it can be recognised directly when the halftone images are covering each other. Author is using a composite method of combining noise- balanced error diffusion (NBEDF) data hiding and kernels-alternated error diffusion (KAEDF) watermarking into one or many error-diffused images.

c) T.-Y. Liu and W.-H. Tsai in, "A New Steganographic method for data hiding in Microsoft Word document by change tracking technique" proposed, the basic idea of technique is to degrade the contents of a cover document D to arrive at another document D' by embedding a secret M message in D during the transformation process. The degradation initiates errors into the degenerated document D' such that the degraded document appears to be a introductory work by a virtual author A'. A stegodocument is then produced from D' by revising D' back to D with the changes being found, making it appear as if author A is correcting the errors in D'. On the other hand, by making use of the change tracking data in the stegodocument , a recipient B of S can easily recover the original document D as well as the degenerated document D' from both of which the embedded information can be extracted. In the embedding place the message bits are embedded using Huffman code alongwith Message Embedding by Text Degeneration and Revision algorithm. The change tracking information comprised of the stegodocument S allows simple recovery of the original document D and the degenerated document D', from both of which the embedded message can be extracted using message extraction algorithm.

d) H. Yang and A. C. Kot in , "Pattern-based data hiding for binary image authentication by connectivity-preserving" proposed, Evaluation of the "flippability" of a pixel to gain good visual quality of the watermarked image. Handling the "uneven embeddability" of the image by embedding the watermark only in "embeddable" blocks. Study of the features in flipping pixels in binary images to attain blind watermark extraction. Exploring different ways of separating the image to achieve larger capacity. Examination on how to locate the "embeddable" pixels in the watermarked image to assimilate cryptographic signature to achieve higher security. A novel blind data hiding scheme based on connectivity-preserving of pixels in a local neighborhood for binary images authentication was proposed. A window of size 3x3 is hired to assess the "flippability" of a pixel in a block. The "uneven embeddability" of the host binary image is handled by embedding the watermark in those "embeddable" blocks based on the three transition criteria. A smaller block size is chosen in order to increase the data hiding capacity.

e) Kede Ma, Weiming Zhang, Xianfeng Zhao, Nenghai Yu, and Fenghua Li in, "Reversible data hiding in encrypted images by Reserving room before encryption" proposed, a novel technique about reserving vacant place before encryption with a RDH algorithm, and hence it becomes easy to hide data reversibly and embed data in the encrypted image. There arises two cases for extraction

1) Case 1: Extracting Data From Encrypted Images: To manage and update personal information of images which are encrypted for protecting clients' privacy, an unsatisfactory database manager may only get access to the data hiding key and have to control data in encrypted domain.

2) Case 2: Extracting Data From Decrypted Images: In Case 1, both embedding and extraction of the data are controlled in encrypted domain. On the other hand, there is a different condition that the user wants to decrypt the image first and extracts the data from the decrypted image whenever needed.

A. Hong Cao, and Alex c. Kot in, "On Establishing Edge Adaptive Grid for Bilevel Image Data Hiding" proposed, a method for authenticating binary host images using an edge-adaptive data hiding method. Uses a simple binary image to show that EAG selects good data carrying pixel locations (DCPL) efficiently in "L -shaped" patterns rather than block-based methods.

The basic idea is to strongly divide a given host image into two regions: the quasi-image region R1 and the flippable region R2 using a data hiding technique. Here R2, can contain a set of separate locations throughout the image. A cryptographic image hash computed based on the region R1 , together with some additional data, e.g., the owner's information, is embedded by altering the region R2 to produce the secure marked image.

3. Various Data Hiding Techniques

a) Binary Matrix, Weight Matrix, XOR Operator Techniques

Binary and Weight Matrix are used as secret keys, whereas The XOR operator is used so that the keys can not be compromised easily.
The inputs to this scheme are as follows.

1) I is a host binary image (i.e., bitmap), which is to be modified to embed data. Here, I is partitioned into blocks of size mxn . Here assume that the size of F is a multiple of mxn.
2) The sender and the receiver share K as a secret key. It is a randomly selected bitmap of size mxn.
3) The sender and the receiver share S as a secret-weight matrix. It is an integer matrix of size mxn whose content satisfie requirements.
4) kr bits which are to be embedded in I contains C as a critical information, where k is the number of mxn blocks in I.

Using all these specifications data was hidden, embedded and extracted and attains following perceptions.

1) Equal Block Size: We use the same block size and compare images quality after data hiding it leads to noisy image.
2) Equal Amount of Embedded Data: The amount of embedded data is further equalized by adjusting the block sizes to compare the image quality after data hiding.

b) Noise-Balanced Error Diffusion (NBEDF) Data Hiding and Kernels-Alternated Error Diffusion (KAEDF) Techniques:

This technique is used to embed watermark in error-diffused images. The visual decoding pattern can be achieved when two or more similar NBEDF images overlap each other. In the improved version of NBEDF, the two halftone images are made from two totally different gray-tone images and it still provides a clear and sharp visual decoding pattern. Printing or other distortions damage the NBEDF methods. Thus, a kernels-alternated error diffusion (KAEDF) technique was proposed. The well-known kernels proposed by Jarvis *et al.* and Stucki were very useful when used in halftone process.

c) Change Tracking Technique:

Steganographic method in which data embedding is used to give the product of a collaborative document authoring effort. In this, the stegodocument seems to be the work of various authors. To provide communication of the authors during the collaborative document authoring process, the word processor records all the exact modifications made by an author and embeds the ways of revision as change tracking information into the document. From this change tracking information, we can get the exact changes made by a previous author, and can recover a previous version of the document when needed.

d) Flippability and embeddability Technique:

This technique works on binary converted images. Flipping the pixels means converting ‘0’ to ‘1’ and vice versa, i.e converting the black pixels to white and white pixels to black. The “flippability” of a pixel or a group of pixels, can be determined by imposing three transition criteria in a 3x3 moving window centered at the pixel. The “embeddability” of a block is not varient in the watermark embedding process, thus the watermark could be extracted without referring the original image. The locations are chosen in such a way that the quality of the watermarked image is guaranteed.

e) Reversible Data Hiding Technique (RDH):

Losslessly vacating room from the encrypted images is quite difficult. Reversing the order of encryption and vacating room, i.e., reserving room previous to image encryption at owner side. It works as follows, the content owner first reserves enough space on host image and then converts the image into its encrypted form using the encryption key. Now, to embedding process data is encrypted in the images reversibly and it needs to occupy data in the vacant places in the image which was priorly emptied. This technique basically comprises of four steps: Generation of encrypted image, data hiding in encrypted image, data extraction and image recovery.

f) Edge-Adaptive Grid Technique:

This technique selects the Data Carrying Pixel Locations (DCPLs) because in flipping technique the selected edge pixels for hiding data can interfere with the local image. This interference is critical for binary document images with corrupted edges by scanning and binarizing noises. Its hides the data in L-shape, since from the studies it is found that the L-shape is very convenient to hide data and this will not hamper the host images quality.

A hybrid method combining the Edge-Adaptive Grid method and Reversible Data Hiding method (RDH) could be used to enhance the quality of the host image i.e with less distortion at the end of the whole process along with more data security and larger payload.

4. Conclusion

Studying these techniques gives an idea of how the methods evolved and merged gradually as per the requirement and the feasibility. More and more stress is given on attaining less noisy host image, less time consumption in order to carry out the whole process, and security to the embedded data into the image. However, previous techniques were less efficient in all such respects, but the recent ones are precisely more dependable and comparatively better.

References

[1] Hong Cao, and Alex c. *Kot*, “On Establishing Edge Adaptive Grid for Bilevel Image Data Hiding” *in* IEEE transaction on information forensic and security, vol. 8, no. 9, september 2013.

[2] Kede Ma, Weiming Zhang, Xianfeng Zhao, Nenghai Yu, and Fenghua Li, “Reversible data hiding in encrypted images by Reserving room before encryption” in IEEE Transactions on information forensics and security, vol. 8, no. 3, march 2013.

[3] H. Yang and A. C. Kot, “Pattern-based data hiding for binary image authentication by connectivity-preserving,” in IEEE trans. multimedia, vol. 9, no. 3, pp. 475–486, apr. 2007.

[4] T.-Y. Liu and W.-H. Tsai, “A New Steganographic method for data hiding in Microsoft Word document by change tracking technique,” in IEEE transactions on information forensics and security, vol. 2, no. 1, march 2007.

[5] S.-C. Pei and J.-M. Guo, “Hybrid pixel-based data hiding and Blockbased Watermarking for error-diffused halftone images”,in IEEE transactions on circuits and systems for video technology, vol. 13, no. 8, august 2003.

[6] Yu-Chee Tseng,Yu-Yuan Chen, and Hsiang-Kuang Pan, “A secure data hiding scheme for binary images” in IEEE transactions on communications, vol. 50, no. 8, august 2002.