A Survey on Different Data Hiding Techniques in Encrypted Images

Sanju Sharma[1] **Saranjeet Singh**[2]
[1]M.Tech. Scholar [2]Assistant Professor
[1,2]Galaxy Global Group of Institutions, Haryana, India

Abstract— In this paper, we are going to have survey on different data hiding techniques and our main focus is on” Reversible data hiding in encrypted images”. In recent year the security of the sensitive data has become of prime and supreme importance and concern. To protect this data or secret information from unauthorized person we use many data hiding techniques like stegnography, cryptography and RDH. In this paper we will discuss on one such data hiding technique called Reversible Data Hiding (RDH). In this instead of embedding data in encrypted images directly, some pixels are estimated before encryption so that additional data can be embedded in the estimating errors. Without the encryption key, one cannot get access to the original image. A RC4 algorithm is applied on the rest pixels of the image and a special encryption scheme is designed to encrypt the estimating errors.Our paper presents a survey on various data hiding techniques and their comparative analysis.

***Keywords**: RDH, stegnography, cryptography*

I. INTRODUCTION

In recent year the security and confidentiality of the data has become prime and supreme importance and concern. To protect this sensitive data from unauthorized person we can use many data hiding techniques such encryption and decryption algorithms encryption is a process of converting the plain text i.e. readable secret information into cipher text. Where decryption is opposite process in this cipher text is converting into plain text. Both sender and receiver of the data can use same key for encrypting and decrypting. But still encryption and decryption have not given 100% security. The reason is that the hacker easily will find the key by capturing the packets and analyzing the cipher text using various software and hardwar. Based on the encryption algorithm and key length the time taken for the hackers to find the key may vary.

Therefore to improve security we can use another method to hide information and send it through network without fear called Stegnography. It is an art of hiding information in multimedia like still images, audio, video over the internet. Basic concept used in stegnography is that any image is made up of pixels. Each pixel represents a color value and depends on the image the pixel size will be from 1 to 4 bytes. These pixels will be stored in computer memory in binary form. Let us consider an image with pixel size 2 bytes. The pixel with 2 byte size can able to represent 216 different colors range from 0000000000000000 to 1111111111111111. Normally for humans eyes the color 1111000011110000 and 1111000011110001 will look like similar because the difference is too low. This means that the change in least significant (LSB) may not be noticed by human eyes. If we alter 1111000011110000 as 0111000011110000 then the color of the pixel will change to another color that is change in MSB will change the color dramatically. This can be identified by everyone. So it is cleared that secret information be stored in LSB and not in MSB of the covered image to reduced the detectable noise. Let us consider a situation that a person named “Ram” wants to send secret information to another person named “Shayam” and the secret information is “Tomorrow meet me in my office”. This information should not be known to a person named “Hari” who is expert in hacking the data or information. In this situation Ram has to take a cover image of size eight times greater than the size of original image or secret information as shown below in figure1.he has to convert the secret information into the binary form and then has to stored the secret information into the LSB of each pixels one by one the resultant image will be sent to Shayam via computer network. Now Hari will catch the packets and he can construct the image send by Ram to Shayam. Now Hari can just see the image and he might think that there is no secret in their communication. This is how one can cheat the hackers using stegnography. The performance of a stegnography can be measured by three factors.These are security, capacity and detectable distortion. Mainly Stegnography involve 4 steps and these are:

(1) Selection of the cover media in which the data will be hidden.
(2) The secret message that is needed to be masked in the cover image.
(3) A function that will be used to hide the data in the cover media and its inverse is to extract the hidden data.
(4) An optional key or the password to authenticate or to hide the data.

In this paper we are representing a survey on different data hiding techniques along with their comparative analysis.The rest of the paper is described as follows: Section 2 presents a survey on different data hiding techniques and their related work. Section 3 presents a comparative analysis of the different data hiding technique.

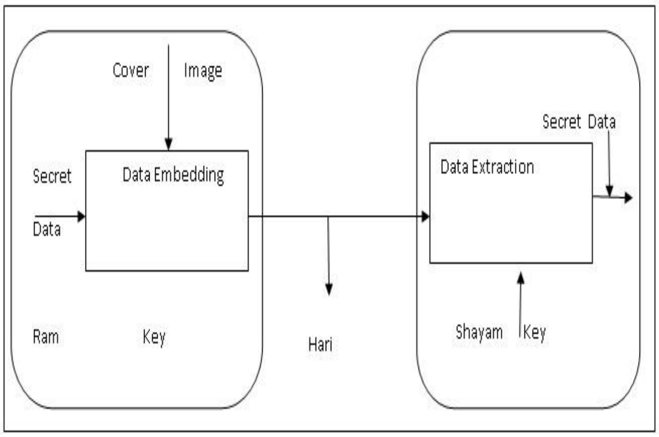

Fig. 1: Steganography

II. DIFFERENT DATA HIDING TECHNIQUES

In this section a survey on various data hiding techniques is presented.

A. *Data Hiding Techniques in Audio Signal*

Kekre et al.proposed two methods to transfer secret data over the network by hiding them in audio signals, thus generating a stego-audio signal.[5,6] In the first method the authors conceal the data in the LSB of the audio by considering the parity of the sample, i.e instead of directly replacing the digitized sample of the audio with secret message, first the parity of the sample is checked and then the secret data is embedded into the LSB. This way it becomes even more difficult for the attacker to guess the bit or data that is being transmitted. In the second approach, XORing of the LSB is performed.The LSB's are XORed and depending on the outcome of this operation and the secret data that is to be implanted the LSB of the data is changed or left unchanged. A different approach is followed by Kondo.Kondo [7] proposed a data hiding algorithm to embed data or information in stereo-audio signals. The algorithm uses polarity of rolling sound which is added to the high frequency signal. In this method the high frequency signal are replaced by one middle channel and then data is embedded. The polarity of rolling sound that is added to each channel is performed to adjust the coherence among these channels. Original signal is not required to extract the hidden data by using this algorithm.

B. *Data Hiding Techniques in Video Sequences*

Li et al. [8] [9] suggested a data hiding technique based on the video sequences. This method implements an adaptive embedding algorithm to select the embed point where the sensitive data is to be hidden. The scheme functions by adopting 4x4 DCT residual blocks and determining a predefined threshold. The blocks are scanned in an inverse zigzag fashion until the first non-zero coefficient is encountered. The value of this coefficient is compared with predefined threshold and if it is greater than the threshold then that pixel is chosen to embed the data.

C. *Data Hiding Techniques Using DNA Sequences*

Abbasy et al. [10] [11] introduced a method to enable secure sharing of resource in cloud computing environments. The proposed method employs the DNA sequences to hide the data. This process also consists of two steps. In the first step DNA sequences is selected and the binary data is converting into this DNA sequences by applying the pairing rules. This step, apart from converting the data also increases complexity by applying the complementary rules and then indexing the garbled sequences. The second method involves the extraction of concealed data from the sequences in where in exactly a reverse operation is performed to the first step.

D. *Reversible Data hiding techniques in encrypted images*

Reversible technique is based on the block division to hide the data in the encrypted image. Instead of embedding data in encrypted image directly some pixels are estimated before encryption so that additional data can be embedded in the estimating error.Rc4 algorithm is applied to the rest pixels of the image. Reversible data hiding has the capability to erase the distortion introduced by embedding step after cover restoration.

RDH techniques have been proposed based on three fundamental strategies such as: Lossless compression appending scheme, difference expansion (DE) and histogram shift (HS).Some recent art are combined the three strategies to residuals of the image such a prediction error s to achieve the better performance. Almost all state of art RDH algorithm consist of two steps .The first step generates a host sequence with small entropy i.e host has a sharp histogram which can be realized by using PE combined with the sorting technique.The second step reversibly embed the message into the host sequence by modifying its histogram with method like HS and DE.Optimal coding methods for modifying histogram proposed by Zhang et al.[12] and Lin and Chung[13].Zhang divided the encrypted image into several block. By flipping 3 Least Significant Bits (LSBs) of a group of specific pixels, one-bit message can be embedded into each block. Hong et al. [14] improved Zhang's method [14] by further exploiting the spatial correlation using a different estimation equation and side match technique. In order to extract data, the two methods rely on decrypted images which maybe unknown for some cases. Aiming for separating data extraction from image decryption, Zhang [15] found the syndromes of a low- density parity check matrix to compress the LSBs of the encrypted image. By doing so, an extra space is created to append additional data. These techniques can only achieve low embedding capacity (achievable largest embedding rate)

- The excellent performance can be achieved in three prospects:
- The proposed method is totally reversible .That is no error occur in data extraction and image recovery steps.
- The PSNR value of decrypted image is much higher than those in previous method.
- Extraction and Decryption are independent, which is more natural and applicable.

Naseem et al. [15] presented an optimized Bit Plane Splicing algorithm to conceal the data in the images. This method used a different approach than the traditional bit plane splicing technique. In this approach instead of just hiding the data pixel by pixel and plane by plane, the process involves hiding the data based on the intensity of the pixels. The intensity of the pixels is described into different range and depending on the intensity of the pixels, the number of bits are chosen that will be used to hide the data in that particular plane. Also, the bits are hidden randomly in the plane instead of hiding them adjacent to each other and the planes are transmitted sporadically thus making it difficult to guess and intercept the data.

Fu.et al. presented some novel method for data hiding in halftone images. The proposed method enables to hide a large amount of data even when the original multitone images are unavailable by forced pair -toggling. The resulting stego-images have high quality and virtually are indistinguishable from the original image

III. COMPARATIVE ANALYSIS OF DIFFERENT DATA HIDING TECHNIQUES

In this section a comparative analysis of different data hiding techniques is presented. The author in [12] [13] have all presented techniques to hide data in still images and reversible data hiding in encrypted images.In this [12] the author embed the data in RGB 24 bit color image by using the linked data structures where in the data hidden data in the image is linked with other data. The advantage of this method is that hiding the randomly then sequential will make it difficult for the hacker to locate it and also without authentication key the hacker will not be able to access the next piece of data in image. Instead of using the whole image as the cover image. The authors in [13] have proposed a method that segments the image into blocks of equal sizes. The process involved in this method is reversible hence there is no loss of hidden data. The approach followed in this method to conceal data is quite different. In this technique the histograms of blocks of images is taken and they are shifted to minimum points of histogram and then data is hidden between these points. The improvement of this technique is that it provides higher capacity to hide the data than the previous method.

The author in [5] [6] use audio signals as the cover media to hide the sensitive data. In [5] the author presents two techniques to hide data. In the first method before hiding the data in LSB of the sample of the audio signal the parity of the sample checked. This method makes the hacker difficult to guess the transmitted data. In the second approach the LSB' are XORed and depending on the result of this operation and the hidden data in the LSB of the sample data is decided to be changed or remain unchanged .In [6]a separate approach is followed which is used in the stereo audio signal to embed data. In the proposed method the polarity of rolling sound is applied to high frequency signal which are then replaced by one middle channel to embed the data.

A Table I is created to provide a brief summary of data hiding techniques with their advantages:

Table 1: A brief summary on various data hiding techniques:

Cover Media	Data hiding Techniques	Advantages
1.Reversible Data Hiding	1. The method Proposed here divided the image into block size s and then uses histogram to embed the data.	1. Rather than sending a single image containing all hidden data, blocks of images can be sent in out of order to confuse the hacker. 2. Since data is hidden in histogram it is difficult to locate the data along with the increase in capacity to conceal the data.
	2. Optimized Bit Plane splicing algorithm where in the pixels are grouped based on their intensity and then the number of bits are to represent the hidden data are chosen.	1.As the bits are grouped based on the intensity of the pixels, more number of darker intensity pixels can be used to represents the hidden data than just the LSB
2. Audio signal	1. In this method author proposed two methods to use the audio signal to hide the data. In the first method, the parity of the sample is checked before replacing the LSB of the sample. In the second approach XORing is carried out.	1. It is difficult to determine the data in the audio signals because the data is not hidden directly in the sample but the polarity is checked before inserting the data.
	2. Secondly the polarity of rolling sound is added to the high frequency channels are used to hide the data.	1. As the polarity of the rolling sound is used to hide the data in the high frequency signals, the stego audio signals generated are more robust and effective during transmission of the information.

IV. Conclusion

In this paper we discussed about various data hiding technique and also presented a notable difference between them. In the introduction section we discussed about various security method like Stegnography, reversible data hiding (RDH) in encrypted image and data hiding in audio signals. In the next section we presented various different techniques to conceal data. Comparative analysis of various data hiding techniques is discussed in last section.

References

[1] Gonzalez, R.C and Woods R.E, Digital Image Processing, Prentice Hall, UpperSaddle River, NJ, 3rd Edition, pp. 385, 2002.

[2] Fridrich, M.Goljan, lossless embedding for all image formats, in: SPIE of Proceedings of Photonics West, Elrctronic imaging and security.vol.4675.San Jose.2002.

[3] M.Celik, G.Sharma, A.Tekalp, E.Saber, Losslesss generalized-LSB IEEE Transaction on image processing (2005).

[4] DM Thodi, J.J Rodriguez, Expansion embedding techniques for reversible watermarking, IEEE Transaction on image processing 16(3) 2007 721-730.

[5] B. Kekre, Archana Athawale, Archana Athawale, Uttara Athawale, "Information Hiding in Applications IJCA, Vol. 7, No. 9, Foundation of Computer Science, New York, USA, pp. 14-19.

[6] B.Santhi, G.Radhika and S. Ruthra Reka, "Information Security using Audio Steganography-A Survey", Research Journal of Applied Sciences, Engineering and Technology, Vol. 4, No. 14, pp. 2255-2258.

[7] K. Kondo, "A Data Hiding Method for Stereo Audio Signals Using the Polarity of the Inter-Channel Decorrelator", IEEE

[8] Yu Li, He-xin, Chen, Yan Zhao,” A New method of data hiding based on H.264 encoded video sequences”, IEEE 10th international conference on signal processing (ICSP),24-28 oct.2010,pp.1833-1836

[9] Xiaoyin Qi,Xiaoni Li, Mianshu Chen, Hexin Chen,”Reasearch on CAVLC audio video synchronization coding approach based on H.264,” IEEE International conference on Uncertainty reasoning and knowledge engineering (URKE),Vol.2,4-7 Aug. 2011,pp.123-126.

[10] Mohammad Reza Abbasy, Bharandharm Shanmugam,” Enabling Data hiding for resource Sharing in Cloud Computing Environment s Based on DNA Sequences “,IEEE World Congress on Services ,4-9 July 2011, pp.385-390