# *Literature Survey on Lossless and Reversible Data Hiding in Encrypted Images with Public Key Cryptography*

[1]Mr. S. R.Divase, [2]Prof A.G.Korke

***Abstract—*** *Data hiding method aims to add some important data in the image .Usually data is hiding in image after encryption of the image with specific algorithm in data hiding section which comes after the encryption section. At Decryption section data is extracted from image and image and data is separated. But there may be chances of loss of original content of image because of data is hiding after the encryption ,means there may be loss of quality, original content of the image. Application area like Military, Medical where minor changes in original content of image affect lots on. This literature survey discusses all the existing data hiding methodology and their performance*.

***General Terms: -*** *Paillier Cryptosystem, Intelligent Crypto Systems.*

***Keyword:-Histogram-Shrink, Difference- Expansion(DE), Plain-Text (PE), Cipher-Text(CT)***

## 1. INTRODUCTION:

Reversible data hiding was mainly proposed for authentication. At starting phase reversible algorithms have small embedding capacity and poor image quality. While the encryption techniques convert plaintext content into ciphertext, the data hiding techniques embed additional data into cover media by introducing small changes. In some distortion-unacceptable cases, data hiding may be performed with a lossless or reversible manner. In number of cases of data hiding, the cover media will results some loss of original content due to data hiding and cannot be extracted back to the original cover media, because some permanent loss has occurred to the cover media even after the hidden data have been retrieved out. In some applications area, such as military, medical diagnosis, it is difficult to recover original content as well as original quality of image. There are different methodology are there of data embedding in reversible or lossless manner as,[1] here pixels with the most used color in a palette image are assigned to some unused color indices for carrying the additional data, and these indices are redirected to the most used color. This way, although the indices of these pixels are altered, the actual colors of the pixels are kept unchanged. On the other hand, we say a data hiding method is reversible if the original cover content can be perfectly recovered from the cover version containing embedded data even though a slight distortion has been introduced in data embedding procedure. A number of mechanisms, such as difference expansion [2], histogram shift [3] and lossless compression [4], have been employed to develop the reversible data hiding techniques for digital images. Recently, several good prediction approaches [5]

## 2. RELATED WORK:

N. A. Saleh, H. N. Boghdad, S. I. Shaheen, A. M. Darwish [1]proposes as **,** Data hiding in encrypted image is one the secure way of transmission of data securely but due to technical aspects after extraction there may be huge loss in quality as well as original content of image. Initially there are problem of capacity of carrying data but recently these problem is overcome there are possibility of huge loss in original quality of image but it can carry high capacity data .Means here we facing only problem of original content of image .To give solution for this problem focusing area is Histogram. Histogram is nothing but graph of image which is drawn on x-axis and y-axis with references to the pixel values .Analysis of histogram is indicates that plotting the graph of original image to avoid the loss of quality of image after decryption issues. Proposed technique works as following manner,

A) Image classification : The embedding technique is depends on the number of the mostly and rarely used colors in the given image. So analysis of image histogram find out color used how many times in that particular given image on which we want to perform operation. So here image are classified according to histogram values , and they calculated on basis of how many times single color is used (that is ripetevely manner) in terms of percentages(%).If any color used number of times it's percentages comes as 60% means it indicates that more than half part of that image having same color. And if any color having percentages 0% it shows that color particular color is not use ,so from this way startup of operation is initiated and performance is measured in the form of percentage and which is shown in table

B) Algorithm :( Based on pixel color count)

i) Start the Process of Histogram means take the image as input

ii) Plot the graph means draws the histogram

iii) Histogram Analysis

a)color representing using values in between [0-255]

b)Count the repetition of particular color

c)Peak value:Color having value 1 is used number of times then count the total number and calculate the peak value eg . 1 used 5000 pixel and it denoted as H[i]=5000 Calculation in percentile as follows formula ***H[i]%= ($N_i$ / N)*100*** Whereas Ni=count of specific used color pixel countN=total number of pixel in that image

iv)End

Sample table0:

Table 1: pixel value calculation table

| Index(color index) | H[i] | H[i](in %) |
|---|---|---|
| 1 | 5000 | 5% |
| 2 | 10000 | 10% |
| . | | |
| . | | |
| 254 | 20000 | 20% |
| 255 | 5000 | 5% |

C) Here unused color can easily find out and in some type of image data embedding capacity not getting as per expectation of the data sender so these are some drawback in this paper.

J. Tian [2] proposed algorithm in the following format and it is in the mathematical form ,

Algorithm:

i)Start
ii) Let two variables x and y
eg x=207 and y=202
Average (l)=((x+y)/2)=((207+202)/2)=409/2=204
Difference((h) =x-y=207-202=5
Consider embedded bit b=1;

iii)Representation
h=5 in binary we can represent as h=$(101)_2$
now new difference (h')=$(101b)_2$ =$(1011)_2$ =11
h'=2*h+b =2*5+1=11
iv)New values
x'= l+[(h'+b)/2]
x'=204+[(11+1)/2]
x'=204+6=210
y'=l-[h'/2] y'=204-[11/2] y'=204-5 y'=199

so (x',y')=(210,199)

now new average(l')=[(x'+y')/2]=[(210+199)/2]=204

l'=204so binary representation is h'=11=$(1011)_2$ =h=$(101)_2$ =5 so Difference expansion mathematical formula is **h'=2*h+b** in above manner difference algorithm works

Zhicheng Ni, Yun-Qing Shi, Nirwan Ansari, and Wei Su[3]Stated as the reversible data hiding approach the basic task of or work of reversible data hiding algorithm is to recover original content as well as quality of the encrypted image. To accomplish this task in this paper they uses the concept of histogram in this they uses zero or minimum point of histogram .This technique having capacity of embed high capacity data into the image but PSNR(peak signal-to-noise ratio) must be greater than 48dB.It uses the images from the CorelDraw database .Where as zero point is nothing but no pixel in given grayscale image and Peak point means maximum number of pixel in given grayscale image

Algorithm: [Embedding on the basis of 1 zero pint and 1 peak point ]

i) Finding the point:-in this step it find out,

zero point :-no pixel in given image

peak point :- maximum number of pixel in image

ii) Phase-I scanning:-

Whole image is scanned in sequential manner and shifting the histogram value by 1 if values in between (154,254) it shifts by 1 that is (155,255)

ii) Phase-II scanning:-

Whole image is scanned in sequential manner and cross checks the grayscale values.So we can insert data from 5kb to 80kb into the range of 512*512*8 values image .Beyond this we cannot increase capacity this limitation of this technique

Mehmet Utku Celik, , Gaurav Sharma,Ahmet Murat Tekalp, Eli Saber [4,]they proposes the data inserting or data

embedding technique based on LSB technique. They tries to recover original content of image without loss but this technique based upon the "host signal" along with this LSB means simply we can say that combination of host signal and the LSB technique.crusial part of this technique is that they needs host signal in form compressed. Algorithm works in following manner

i)Embedding:

Here generation of the water marked image which is combination of host signal and message data.

ii)Extraction:-

In extraction it takes water marked image as input and extract the original host signal and embedded data..

Algorithm LSB:

i) Embedding one bit:

LSB of each signal sample is over written by a payload data bit embedding one bit of data per input sample

ii) Embedding two bit:

If application required additional two or more bit over written per sample

iii)Extraction

During extraction scanning order is same as embedding and payload data is reconstructed

Xiaocheng Hu, Weiming Zhang, Xiaolong Li, and Nenghai[5] Yu proposes prediction –error expansion(PEE) based reversible data hiding technique. it works in two steps as

a) Histogram Generation:-in this step histogram is generated using pixel prediction methodology.

b) Message Embedding:- in this step secret message is inserted into the prediction-error by expanding and shifting histogram.

Algorithm:

i) Image divide:

Here image divide into the square blocks and classify into the different class

ii) Apply K-means:- Apply K-means clustering algorithm for the creation of cluster

iii) Estimate:-

Estimate square prediction

iv) Threshold value:-

Select threshold value from region which contains group of cluster

v)Classify each image block into the class

vi)Shape parameter:-

select optimize shape parameter

vii)Prediction:-

predict image pixel using estimated predicators

viii)Encoding:-

Encodes all predictor coefficient and then record coded overhead bit stream for the LSB

Xinpeng Zhang [ 6] In this paper they proposes the optimal rule by maximizing iterative algorithm

Algorithm:

1) Histogram Creation:

Represent the histogram of data in following format

$H=\{\ldots h_{-3h},h_{-2},h_{-1},h_0,h_{+1},h_{+2},h_{+3}\ldots\}$ where $h_i$ available of data with value i.

2) Create Transfer Matrix:

Original value i and new value k represented in following matrix form $t_{k,j}$

$$\begin{bmatrix} tM1,M2\ldots. & tM1,M2 \\ . & . \\ tM2,M1 & tM2,M2 \end{bmatrix}$$

3) Calculation of vital factor

a)Histogram values

b) Entropy function

c)Apply algorithm: Iterative method

i)Initialization: matrix initialization and distortion level(D=0) and pure payload(P=0)

ii)Calculate updated histogram value H'

iii)Check entropy function

iv) Assign new value:assign new values of D and P and go to step ii

Weiming Zhang, Xiaocheng Hu, Xiaolong Li, and Yu Nenghai[7] uses RCC(recursive code construction) and the rate-distortion bound (RDB)of reversible data hiding for estimation of rate distortion bound (RDB) and for the excutation of RCC, one should first estimate the optimal transition probability matrix (OTPM).But OTPM cannot useful in all cases it not works in some cases. So here they proposes unified framework bating OTPM for all type of Applications.

This method works in following two steps

a)Optimal Transition Probability With NCE Property
b)Recursive Code Construction

Lian, Zhongxuan Liu, Zhen Ren, and Haila Wang[8] proposes the methodology of commutative encryption and watermarking in video compress this method works in following manner
A) Encryption Process:

1) Motion information (MVD) Encryption:-it used for encryption of object which having property of motion it means it used in video

2) Texture information (IPM) Encryption:-it useful in encryption of object which having static property means it useful in text encryption

B) Watermarking Process:

1) Block Selection*:* Only the Luma blocks satisfying the following conditions are watermarked.

i) The residue block is nonzero.

ii) For I/P-frame, the residue DCT block is composed of onlyac's.

iii) For B-frame, the residue DCT block is composed of either dc's or ac's.

2) Coefficient Selection

3) Watermark Embedding

M. Cancellaro , F.Battisti , M.Carli , G.Boato , F.G.B.DeNatale , A.Neri[9]proposes the specific methodology for commutative digital image watermarking and encryption on the basis of Haar transform. In the era of digital computer basic things are communication and security for that communication or data security is very important pillar. Here they basically stick to layered architecture for cryptography and watermarking. Methodology works in following manner

Algorithm:

1)Take input digital image (D)
2)Let
$f_v$: it is used to hide watermark V into D
k:encryption key
$f_{ch}$:Cipher of original content
$Dv= f_v(D,k)$
$De= f_{ch}(D,k)$
3)For commutative performance following conditions are necessary
$Dv,e= f_v(f_{ch}(D,k)v) = f_{ch}(f_v(D,v)k)$

Xinpeng Zhang [10] proposes reversible data hiding in the encrypted image in this paper they focused on data hiding after the encryption of image not before the encryption

Algorithm:

1)Image Encryption:-Original content or image must not be in uncompressed format with encryption key operation of encryption is performed and result is cipher text

2) Data Embedding:-Now the main task is that inserting message or data which we want to send securely to receiver domain by using technique of dividing each block, pseudo-randomly divide into two sets of block and operation is carried out

3)Data Extraction and Image Recovery:-At the receiver side main goal is extraction of data and content recovery means without change in quality of image that is receiver want original image as it is but due to some change during data embedding there are chances of small change in original image which is not acceptable

Table 2 : Comparison table

| Sr.No | Name of Paper | Technique | Drawback |
|---|---|---|---|
| 1 | High capacity lossless data embedding technique for palette images based on histogram analysis | Lossless Compression Method | Image quality is reduced due to compression in pixel |
| 2 | Reversible Data Embedding Using a Difference | Difference expansion | Distortion of the image |

| | | | |
|---|---|---|---|
| | Expansion | | |
| 3 | Reversible Data Hiding | RDH | Low capacity (capacity is not enough to hide high capacity payload) |
| 4 | Lossless Generalized-LSB Data Embedding | LSB Modification | Single bit plane in small images does not offer space for hiding hash after comparison, so two or more bit-planes required and artifacts must visible if offset on security |
| 5 | Minimum Rate Prediction and Optimized Histograms Modification for Reversible Data Hiding | Prediction-error expansion | Low capacity (capacity is not enough to hide high capacity payload) |
| 6 | Reversible Data Hiding With Optimal Value Transfer | Optimal Value Transfer | Time complexity is high as compared to other techniques |
| 7 | Optimal Transition Probability of Reversible Data Hiding for General Distortion Metrics and Its Applications | Recursive Code Construction | Not reaches to rate distortion bound |
| 8 | Commutative Encryption and Watermarking in Video Compression | Haar Transform | More expensive to calculate Haar Wavelets |
| 9 | A commutative digital image watermarking and encryption method in the tree structured Haar transform domain | Haar Transform | More expensive to calculate Haar Wavelets |
| 10 | Reversible Data Hiding in Encrypted Image | RDH | Loss the quality of original content |

**3 Conclusion:** This paper compares various techniques and methodology and it's pit-falls which used for data hiding in the encrypted image. Our main goal is not only hiding importance data in encrypted image but also extraction of data without small change in the original content or original image. Original content recovery plays vital role in the secure communication and gives best result in field like medical, army.

ACKNOWLEDGMENT:

I would like to take this opportunity to acknowledge the contribution of certain people without which it would not have been possible to complete this paper work. I would like to express my special appreciation and thanks to my advisor Professor Ashok Korke, you have been a overwhelming mentor for me.

REFERENCES:

[1] N. A. Saleh, H. N. Boghdad, S. I. Shaheen, A. M. Darwish, "High Capacity Lossless Data Embedding Technique for Palette Images Based on Histogram Analysis," *Digital Signal Processing*, 20, pp. 1629−1636, 2010.

[2] J. Tian, "Reversible Data Embedding Using a Difference Expansion," *IEEE Trans. on Circuits and Systems for Video Technology*, 13(8), pp. 890−896, 2003.

[3] Z. Ni, Y.-Q. Shi, N. Ansari, and W. Su, "Reversible Data Hiding," *IEEE Trans. on Circuits and Systems for Video Technology*, 16(3), pp. 354−362, 2006.

[4] M. U. Celik, G. Sharma, A. M. Tekalp, and E. Saber, "Lossless Generalized-LSB Data Embedding," *IEEE Trans. on Image Processing*, 14(2), pp. 253–266, 2005.

[5] X. Hu, W. Zhang, X. Li, and N. Yu, "Minimum Rate Prediction and Optimized Histograms Modification for Reversible Data Hiding," *IEEE Trans. on Information Forensics and Security*, 10(3), pp. 653-664, 2015.

[6] X. Zhang, "Reversible Data Hiding with Optimal Value Transfer," *IEEE Trans. on Multimedia*, 15(2), 316−325, 2013.

[7] W. Zhang, X. Hu, X. Li, and N. Yu, "Optimal Transition Probability of Reversible Data Hiding for General Distortion Metrics and Its Applications," *IEEE Trans. on Image Processing*, 24(1), pp. 294-304, 2015

[8] S. Lian, Z. Liu, Z. Ren, and H. Wang, "Commutative Encryption and Watermarking in Video Compression," *IEEE Trans. on Circuits and Systems for Video Technology*, 17(6), pp. 774−778, 2007.

[9] M. Cancellaro, F. Battisti, M. Carli, G. Boato, F. G. B. Natale, and A. Neri, "A Commutative Digital Image Watermarking and Encryption Method in the Tree Structured Haar Transform Domain," *Signal Processing: Image Communication*, 26(1), pp. 1−12, 2011.

[10] X. Zhang, "Reversible Data Hiding in Encrypted Image," *IEEE Signal Processing Letters*, 18(4), pp. 255−258, 2011.